

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 20:41:32 ; Search time 8423 Seconds  
(without alignments)  
10993.468 Million cell updates/sec

Title: US-10-731-525-7  
Perfect score: 1629  
Sequence: 1 gcacggaggagaccata.....taaaaaaaaaaaaaaaaaa 1629

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues  
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_in.\*  
3: gb\_env.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pr.\*  
9: gb\_ro.\*  
10: gb\_sts.\*  
11: gb\_sy.\*  
12: gb\_un.\*  
13: gb\_vi.\*  
14: gb\_htg.\*  
15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1629	100.0	1629	6 AR533623	AF533623 Sequence
2	916.8	56.3	1705	15 AF452450	AF452450 Glycine m
3	904	55.5	1795	6 AR533622	AF533622 Sequence
4	718.4	44.1	1648	15 AY353089	AY353089 Populus x
5	706.8	43.4	1753	15 CRU63784	U63784 Catharanthu
6	664.8	40.8	1835	15 BT013254	BT013254 Lycopersi
7	663.8	40.7	1386	15 AY568717	AY568717 Lycopersi
8	663.6	40.7	1608	15 SCJ001208	AJ001208 Brassica
9	659.4	40.5	103129	15 AP004494	AP004494 Lotus cor
10	657.2	40.3	1377	15 AY103313	AY103313 Arabidops
11	657.2	40.3	1555	15 AY093319	AY093319 Arabidops
12	657.2	40.3	1654	15 AY062665	AY062665 Arabidops
13	657.2	40.3	1702	15 AY087591	AY087591 Arabidops
14	657.2	40.3	1709	15 AF428445	AF428445 Arabidops
15	657.2	40.3	1712	15 AY054175	AY054175 Arabidops
16	656.6	40.3	1710	15 AF424582	AF424582 Arabidops
17	655	40.2	1714	15 AY088011	AY088011 Arabidops
18	653.2	40.1	1398	15 BT002612	BT002612 Arabidops

19	652.4	40.0	1763	15 ATU53865	US3865 Arabidopsis
20	651.6	40.0	1398	6 CQ805882	CQ805882 Sequence
21	651.6	40.0	1398	6 AX412862	AX412862 Sequence
22	651.6	40.0	1398	6 AX505803	AX505803 Sequence
23	651.6	40.0	1398	6 AX651565	AX651565 Sequence
24	650.2	39.9	1748	15 ATU53864	US3864 Arabidopsis
25	620.8	38.1	1482	15 ATU56922	US6922 Arabidopsis
26	615.8	37.8	1634	15 ATU43412	U43412 Arabidopsis
27	602.8	37.0	73519	14 AP007975	AP007975 Lotus cor
28	594.2	36.5	1700	15 SCJ001207	AJ001207 Brassica
29	588.2	36.1	1687	15 AY088665	AY088665 Arabidops
30	586.6	36.0	1712	15 ATU96045	U96045 Arabidopsis
31	578.4	35.5	1783	15 AF360192	AF360192 Arabidops
32	577.4	35.4	1365	6 CQ805458	CQ805458 Sequence
33	577.4	35.4	1365	6 AX412892	AX412892 Sequence
34	577.4	35.4	1365	6 AX412893	AX412893 Sequence
35	577.4	35.4	1365	6 AX507610	AX507610 Sequence
36	576.6	35.4	1385	15 AY040005	AY040005 Arabidops
37	576.4	35.4	1365	15 AF023167	AF023167 Arabidops
38	570.6	35.0	1738	15 ATU53866	US3866 Arabidopsis
39	568	34.9	1945	15 AF016282	AF016282 Arabidops
40	568	34.9	47319	6 AX059533	AX059533 Sequence
41	568	34.9	103632	15 AF074021	AF074021 Arabidops
42	568	34.9	200001	15 ATCHRIV13	AL161501 Arabidops
43	562.4	34.5	1450	15 ATU56921	US6921 Arabidopsis
44	561.6	34.5	2945	15 AF016284	AF016284 Arabidops
45	561.6	34.5	103270	15 ATFIN20	AL022140 Arabidops

ALIGNMENTS

RESULT 1  
AR533623  
LOCUS AR533623 1629 bp DNA linear PAT 08-OCT-2004  
DEFINITION Sequence 7 from patent US 6730827.  
ACCESSION AR533623  
VERSION AR533623.1 GI:53923592  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1629)  
AUTHORS Falco,S.C.  
TITLE Genes encoding plant adenosine 5'-phosphosulfate reductase  
JOURNAL Patent: US 6730827-A 7 04-MAY-2004;  
E. I. du Pont de Nemours and Company; Wilmington, DE  
FEATURES  
source  
1. .1629  
/organism="unknown"  
/mol\_type="genomic DNA"

Query Match	100.0%	Score 1629;	DB 6;	Length 1629;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1629;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GCACGAGGAGAGAACCCATACAGCTAGTTAATGGCCCTCGCTTCACTTCTTCAATTTC	60	
Db	1	GCACGAGGAGAGAACCCATACAGCTAGTTAATGGCCCTCGCTTCACTTCTTCAATTTC	60	
Qy	61	CGCACCACCACTTCCACCTTCCCATCGGAAACCCAACTTCCGCAATTTGGGTCAATTAG	120	
Db	61	CGCACCACCACTTCCACCTTCCCATCGGAAACCCAACTTCCGCAATTTGGGTCAATTAG	120	
Qy	121	GATTTCGGAGAGGCCCATTCGGAGCGCGCTTAATTTCAATTTATCTCAAGACGGAGCTT	180	
Db	121	GATTTCGGAGAGGCCCATTCGGAGCGCGCTTAATTTCAATTTATCTCAAGACGGAGCTT	180	
Qy	181	GGTAAAGCCGGTTAAACGCGAACTCCACGCAAGGATTCCTCTCGCAGCAAC	240	
Db	181	GGTAAAGCCGGTTAAACGCGAACTCCACGCAAGGATTCCTCTCTCGCAGCAAC	240	



```
Qy 41 GCTTCACTTCTTCAATTTCCGCAACCAACTTCCACCTTCCATCATCGGAACCCAAACTT 100
Db 55 GCTCTGCGAGCTTCTAGCTCTTCTTCTCTCGTCTCGGATCTTCTCTCGAGCGCAAAAGCT 114
Qy 101 CCGCAAAATGCGGTCAATTTAGGATTTCCGAGAGGCCATTTGAGGCGCGCTTAATTTCAAT 160
Db 115 CCTCAAAATCGGTTCGTTTCGTTTCCGAGAGGTCGCTTGTTCGTCTGTTGTGTCAT 174
Qy 161 TTATCTCAAGACGCGAGCTTGGTAAGCCGTTTAAACGCGCAACCTCCACGCAAGGATTC 220
Db 175 GTAACCTCAAGACGCTCTCTGCTGAGGCCACTCAACGCGCAACCGCAACGGAATCT 234
Qy 221 ATTGTTCTCTCGCAGCAACCAACCATGTTGCTTCTGCTTCTGAGAGCAAGAGAGAT 280
Db 235 ATTGTTCTCTCTGAG---CAACTATCGTCTCTGAGGTTGGAAGGAGAGAGAGAT 291
Qy 281 TTTGAAACAGATAGCGAGTGTCTCGAAATGCTTTTCACTCTTGAATCATGATAGAGCC 340
Db 292 TTTGAGCAATAGCGAAGACCTTGAANAATGCAATCTCTCTTGAANAATGATAGAGGCC 351
Qy 341 CTCGACAAATTCGGCAACGACATAGCTATTGCCCTTCAAGTGTCTGAAGATGTTGCTTTG 400
Db 352 CTCGAGAAATTTGGAAACGACATCGCTATTGCAATTTAGTGTGCTGAAGATGTTGCTTTG 411
Qy 401 ATTGAGTATGCGAAATTTGACGGGTGCAACCTTTTAGGGTTTTCACTTTGGAACATGCGAGA 460
Db 412 ATTGAGTATGCAATTTTGAACGGGTGCAACCTTACAGAGTGTGTTAGTCTTGAACATGCGAGA 471
Qy 461 CTGAACCCAGAACTTATCACTTTTGTGATCGGTTGAGAGCAATTTAGNAATTCGCAAT 520
Db 472 CTGAATCCAGAACTTCAAAATTTTGTGCGGTTGAGAGCAATTTAGNAATTCACAT 531
Qy 521 GAGTACATGTTTCCCTGATGCTGTTGAGGTTTCAAGCATTTGAGAGTAAAGGGGTTATTC 580
Db 532 GAGTACATGTTTCCCTGATGCGTTGAGGTTTCAAGCATTTAGTAAAGGCTCTTC 591
Qy 581 TCTTTCTACAGAGTATGGCAACCAAGATGTTTGAAGGTGAGAAAGGTGAGCCCTTTAAG 640
Db 592 TCATTTTACGAGGATGGCATCAAGAGTGTGCGGAGTGAAGAGGTGAGGCCCTTGAG 651
Qy 641 AGGSCCTTAAGGCTCTCAGAGCATGATAACTGTTGAGAGGAAAGACCATCCTGCT 700
Db 652 AGAGCCCTTAAGGCTCTCAAGCATGATGATCACTGACAGAGAAAGATCATGCTCTCTG 711
Qy 701 ACTAGGCTGAAATACCGGTTGTTTCAAGTGTGATCCGCTTTTGAAGGAATGGATGGGA 760
Db 712 ACTAGGCTGAAATCCCTGTTGTCAGGTTGATCTGTTTGAAGGACTGGATGGTGA 771
Qy 761 ATTGGAAGCTTGTGAAGTGAACCCCTGTTGCAAAATGTGAAGGCCCATGACATATGAAC 820
Db 772 ATTGGCAGCCTGGTGAAGTGAACCCGTTGCAAAATGTTAATGCTTGAACATATGGAGC 831
Qy 821 TTCTTTAGGACCATGATGTCCTGCAATTCCTTGCATGCAAAAGATATGTTCCAT 880
Db 832 TTCTTTAGGACCATGATGTTCTGTAATTCATTTGCAATTTCCCAAGGATATGTTTCAT 891
Qy 881 GGGTGTGAGCCCTGCACTAGGCTGTTTATACCTGGGCAACATGAAGGGAAGGAGGTGG 940
Db 892 GGGTGTGAGCCGTGCAACAGGCCAGTTTACCGGCTCAACATGAAGAGAGAGAGGTGG 951
Qy 941 TGGTGGGAGATGCCAAGCTAAGGAATGTTGCTTTCACAAAGGAATGTAAAGCAGCAG 1000
Db 952 TGGTGGGAGATGCCAAGCCAGGAGTGTGCTTTCACAAAGGTAATATAAAACATGAA 1011
Qy 1001 AAAGAGAGGATGTTAATGGAATGGGCTATCCCAATCCCATGCAATGTTGATGCTACC 1060
Db 1012 GATGCTGCCAGCTTAATGGAATGGGGCTTCCCAAGCAAAATGGA-----TCTGCC 1062
Qy 1061 ACTGTGCTGACATTTTCAACAGCCGAATGTAGTTAACTTTGAGCAGGACTGGAATTTAG 1120
Db 1063 ACTGTGCTGACATTTTCACTCCAGGATGTGTTGAGCTTGAAGAGGCTCTGGGATTTAG 1122
Qy 1121 AATTTGGCAAAATTTGGAGGACCGAAAGGAACCATGGCTTGTGCTTTATGCAACCATGG 1180
```

```
Db 1123 AATTTGGCAAAATTTGGAGAACCGAAAGAACCCCTGGCTTGTGTGCTCTATGCAACATGG 1182
Qy 1181 TGCCCTTACTGCCAGGCTATGGAGGAATCTTATGCTTACCTTAGCAGACAGCTTAGCAGG 1240
Db 1183 TGCCGCTTCTGTCAAGGCTATGGAGGAATCATATGTTGATTTGGCAGAAAAGTTAGCTGG 1242
Qy 1241 TCACAGGAGTGAAGGTTGGAAAATTTAGAGCAGATGGAGAACAGAAAGAAATTTGCAAG 1300
Db 1243 T---CAGGAGTGAAGGTTGCAAAATTCAGAGCGATGGAGACAGAGGAATATGCAAG 1299
Qy 1301 AGTGAATGCAATTTGGGAAGCTTCCCTACGATATTTATTTTCCAAAGCATTCGCTCGG 1360
Db 1300 ACTGAATGCAATTTGGGAAGCTTCCCAACAATACTACTCTTCCCAACAACCTTCT--- 1356
Qy 1361 CCAACAATAAAGTATCCCTCAGAAAAGAGAGATGTTGATCTCTTGTGATGGCATTTGTAAT 1420
Db 1357 CAACCAATTAAGTACCTTCGAAAAGAGAGATGTTGATTCATTTGACGGCATTCGTGA 1416
Qy 1421 GCCTTAAGATCAGGAT 1436
Db 1417 GCTTTACGGTGATGGT 1432

RESULT 3
AR533622
LOCUS AR533622 1795 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 5 from patent US 6730827.
ACCESSION AR533622
VERSION AR533622.1 GI:53923591
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1795)
AUTHORS Falco, S.C.
TITLE Genes encoding plant adenosine 5'-phosphosulfate reductase
JOURNAL Patent: US 6730827-A 5 04-MAY-2004;
E. I. du Pont de Nemours and Company; Wilmington, DE
FEATURES
source 1..1795
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 55.5%; Score 904; DB 6; Length 1795;
Best Local Similarity 81.3%; Pred. No. 2e-227;
Matches 1103; Conservative 0; Mismatches 235; Indels 18; Gaps 4;
Qy 81 CATCATCGGAACCCAAACTTCCGCAAAATTTGGTCAATTTAGGATTTCCGAGAGGCCCATTTG 140
Db 135 CTTTCATCGGACGCTAAAGCTCCGCAAAATTTGGTTCCTTTCGTTTCCGAGAGGCCCTCAAG 194
Qy 141 GAGGCGCGTTAATTTCAATTTATCTCAAAAGACGAGCTTGGTAAAGCCGCTTAAACGCG 200
Db 195 TTTTCGCTCGTGTGTTTAAATTTAACTCAAAAGACGCTCTCGGTGAGGCCCATCAATGCG 254
Qy 201 AACCTCCAGCAGAGGATTCATTTGTTCTCTCCGAGCAACCAATGCTGTTGCTTCTGCTT 260
Db 255 AACCGCAACGGAATGATTTCTGTTTCTCTCTTGAG---CAACTATGTTGCTCTGAGG 311
Qy 261 CTGAGACGAAAGAGGAAGATTTTGAACAGATAGCAGTATCTCGACAATGTTTCACCTC 320
Db 312 TTGAGAGGAGAAAGAAAGATTTTGAGCAATTTAGGAAGACCTTGAATAATTCATCTCTC 371
Qy 321 TTGAAATCATGATAGAGCCCTCGACAAATTCGGCAACGACATAGCTATTGCTTCAAGT 380
Db 372 TTGAGATTTAGGATTAAGGCCCTCGAGAAATTTGGGAACGACATCGCTATTGCTTTAGT 431
Qy 381 GTCTGAGATGTTGCTTTGATTTGATGATGCGAAATTTGACGGTTCGACCCCTTAGGGTTT 440
Db 432 GTCTGAGATGTTGCTTTGATTTGATGATGATGCAATTTGACGGGTCGACCCCTACAGAGT 491
```

```
Qy 441 TCAGTTTGGACACTGGGAGACTGAACCCAGAACTTATCAACTTTTGTGCGGTGAGA 500
Db 492 TTAGTCTTGACACTGGGAGACTGAACCCAGAACTTATCAAAATTTTGTGACGCTGTGAGA 551
Qy 501 AGCATTATGGAATTCGCATTGAGTACATGTTTCCCTGATGCTGTGAGGTTTCAGGCATTGG 560
Db 552 AGCATTATGGAATTCATTGAGTACATGTTTCCCTGATGCGTGTGAGGTTTCAGGCATTAG 611
Qy 561 TGAGGAGTAAGGGGTTATCTCTTTCTACGAGGATGGCCACCAAGAGTGTGTGACGGTGA 620
Db 612 TAAGAACTAAGGGGCTCTTCTCATTTTACGAGGATGGCCATCAAGAGTGTGTGAGAGTAA 671
Qy 621 GAAAGGTGAGCCCTTAAAGGAGCCCTTAAAGGCTCTCAAGGATGATTAAGTGTGAGTGA 680
Db 672 GAAAGGTGAGCCCTTGAAGGAGCCCTTAAAGGCTCTCAAGGATGATTAAGTGTGAGTGA 731
Qy 681 GGAAGACCACTGACCTGCTACTAGTCTGAAATACCGGTTTGTTCAGGTTGATCCGGCTT 740
Db 732 GAAAGACCACTGCTCTGCTACTAGTCTGAAATACCGGTTTGTTCAGGTTGATCCGGCTT 791
Qy 741 TTGAGGGAATGGATGGTGGAAATCGAACTTGGTGAAGTGGAAACCCCTGTTGCAAAATGTA 800
Db 792 TTGAGGGAATGGATGGTGGAAATCGAACTTGGTGAAGTGGAAACCCCTGTTGCAAAATGTA 851
Qy 801 AGGCCATGACATATGGAATCTTCTTAGGACCATGAATGCTGCTGAATTCCTTGCATG 860
Db 852 ATGCTCTAGACATATGGAATCTTCTTAGGACCATGAATGCTGCTGAATTCCTTGCATG 911
Qy 861 CAAAGGATATGTTTCCATTCGGTGTGAGCCCTGCACCTAGGCTGTTTACCTGGGCAAC 920
Db 912 CCAAGGATATGTTTCCATTCGGTGTGAGCCCTGCACCTAGGCTGTTTACCTGGGCAAC 971
Qy 921 ATGAAAGGGAAGGAGGTGGTGGGAGGATGCCAAAGCTTAAAGAAATGTGGTCTTCA 980
Db 972 ATGAAAGGGAAGGAGGTGGTGGGAGGATGCCAAAGCTTAAAGAAATGTGGTCTTCA 1031
Qy 981 AAGGAATGTAAAGCAGAGAGGAGGATGTTATGGAATGGGCTATCCCAATCCC 1040
Db 1032 AAGGTAATTTGAAACAGGAGAGATGTCGCCAGCTTAAATGGAATGGGACCTCCCAAGGA 1091
Qy 1041 ATGCAATATGTTGATGCTACCTGCTGCTGACATTTTCAACAGCCCGGAATGTAGTAACT 1100
Db 1092 A-----TGGCTCTGCCACTGTGCTGACATTTTCACTCCAGAAATGTGTGAGCT 1142
Qy 1101 TGAGCAGGACTGGAATTTGAGAAATTTGGCAAAATTTGGAGGACCGAAAGAAACATGGCTG 1160
Db 1143 TGAGCAGGCTCGGATTTGAGAAATTTGGCAAAATTTAGAGAACCGAAAGAACACATGGCTG 1202
Qy 1161 TTGTGCTTTATGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1220
Db 1203 TTGTGCTCTATGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1262
Qy 1221 TAGCAGACAAATTTAGCAGGCTCAACAGGAGTGAAGTTTGGAAATTTAGAGCAGATGAG 1280
Db 1263 TGCCAGAGAGTTAGCAGGCT---CAGAGTGAAGTTTGGAAATTTAGAGCAGATGAG 1319
Qy 1281 AACGAAAGAAATTTGCAAGAGTGAATGCAATTTGGGAGGCTTCCCTACGATATATTTT 1340
Db 1320 AGCAGAGGAATATGCAAGAGTGAATGCAATTTGGGAGGCTTCCCTACGATATATTTT 1379
Qy 1341 TCCCAAGCATGCTGCTGCGGCAACAAATAAGTATCCCTCAGAAAGAGAGATGTTGATT 1400
Db 1380 TCCCAAGCATGCTGCTGCGGCAACAAATAAGTATCCCTCAGAAAGAGAGATGTTGATT 1436
Qy 1401 CTTTCATGCGCATTTCTGATGCTTAAATGCTTAAAGATGAGAT 1436
Db 1437 CATTGACGCGCATTTCTGATGCTTAAATGCTTAAAGATGAGAT 1472
```

RESULT 4  
AY353089 1648 bp mRNA linear PLN 12-JAN-2005  
LOCUS  
DEFINITION Populus x canescens adenosine 5' phosphosulfate reductase mRNA,

complete cds.  
AY353089 GI:34099826  
AY353089.1  
Populus alba x Populus tremula  
Populus alba x Populus tremula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids I; Malpighiales; Salicaceae; Populus.  
1 (bases 1 to 1648)  
Kopriva,S., Hartmann,T., Massaro,G., Hoenicke,P. and Rennenberg,H.  
Regulation of sulfate assimilation by nitrogen and sulfur nutrition  
in poplar trees  
Trees (Berl. West) 18, 320-326 (2004)  
2 (bases 1 to 1648)  
Hartmann,T., Hoenicke,P., Wirtz,M., Hell,R., Rennenberg,H. and  
Kopriva,S.  
Regulation of sulphate assimilation by glutathione in poplars  
(Populus tremula x P. alba) of wild type and overexpressing  
gamma-glutamylcysteine synthetase in the cytosol  
J. Exp. Bot. 55 (398), 837-845 (2004)  
3 (bases 1 to 1648)  
Kopriva,S.  
Direct Submission  
Submitted (28-JUL-2003) Institute of Tree Physiology,  
Georges-Koehler-Allee 053, Freiburg 79085, Germany  
Location/Qualifiers  
1. .1648  
/organism="Populus alba x Populus tremula"  
/mol\_type="mRNA"  
/db\_xref="taxon:80863"  
/clones="INRA 717 1B4"  
74. .1471  
/codon\_start=1  
/product="adenosine 5' phosphosulfate reductase"  
/protein\_id="AAO57202.1"  
/db\_xref="GI:34099826"  
/translation="MALVSSSSSISASGFSRSSQELKAPQFGSKLLDRPTTFVNV  
SQRCAVKPLNAPKRNQSVVPLAATAEIAEKVEEDYEKLAKELVNGSPLEIM  
DKALEYDDIAIAFGAEVALIEYAKLTGRPFVFLSDTLGRNLNPTTFHFDQVEKH  
YGRLEYMFDPAVEQALVRNKLGFYFEDGHQCCRVKRVPLRLKALGLRAWITGQ  
RKQSPCTSEIPVQVDPVPEGLDGGAGSLIKWNPANVEGQDVWKFPLRTMDVPVNS  
LHSGYISIGCEPCTRPVLQOHEREGRWMDATAKECGHLKNGDAOLNGNG  
NGAHRANGATVADIFNSNLNLSRGIEINLLKLRNKEPWLVLVAPWQFCQGM  
ASYVELADKLKAGSVKVGKFRADGDKQFSGKQELQLGFFLILFPKHSRPIKYPSE  
KRDVDSLMTFVNLRL"

## ORIGIN

Query Match 44.1%; Score 718.4; DB 15; Length 1648;  
Best Local Similarity 72.3%; Pred. No. 2.1e-178; Indels 21; Gaps 5;  
Matches 1009; Conservative 0; Mismatches 366;  
Qy 43 TTTCACTCTTCAATTTCCGCAACCACTTCCACCTTCCCATCATCGGAACCAAACTTCC 102  
Db 91 TCTCTCTCTCGATTCTGCTTCTGGGTTTTCAGGTTCTAGTCAAGAACTTAAAGCACC 150  
Qy 103 GCAAATTTGGGTCAATTAGGATTTGGAGAGGCCCATTTGGAGCGCGCTTAATTTCAA 162  
Db 151 CCAGTTTGGTCTTTTAACTGTTAGATCGGCCGA-----CAACATTTGTTAATGTTTC 204  
Qy 163 ATCTCAAGAGCGAGCTTGGTAAAGCCCGTTAAACCGCAACCTCCACCAAGGATTCAT 222  
Db 205 TTCAAAAGACGCTGTGCTGTAAAGCCCTTTAAATGCTGAGCCTTAAAGAAATGTTCTGT 264  
Qy 223 TGTTCCTCTCGCAGCAACCAACCATCGTTGCTTCTGTTCTGAGACGAAA-----GAGA 276  
Db 265 TGTTCCTCTCGTCAACTATCGTGTCTCCAGAGATTCAGAGAAAGTAGAGGTGAGGA 324  
Qy 277 AGATTTTGAACAGATAGCAGTGTATCTCGACATGCTTCACTTGAATCATGGATAG 336  
Db 325 AGATTTTGAAGATTTAGCTAAGGAGCTTGTAAATGGTTCTCTCTTGGAAATTTATGATAA 384



```
Qy 337 AGCCTCGACAAATTCGGCAACGACATAGCTATTATGCTTCAAGTGGTCTGAAGATGTGC 396
Db 385 GGCATTGAGAAATTTGGGATGACATTTGCTATTGCTTTTCAAGTGGTCCGCAAGATGTGC 444
Qy 397 TTTGATTGAGTATGCGAAATTCAGCGGTTCGACCTTTAGGGTTTTCAGTTTGGACATGG 456
Db 445 TTTGATTGAGTATGCAAGTTGACTGGCGGACCGTTTCAAGGTTTCTCATTGGAATCGGG 504
Qy 457 GAGACTGAACCCAGAACTTATCAACTTTTGTATGCGGTTGAGAGCAATTTATGGAATTCG 516
Db 505 GCGTGTGAACCCGAGACATACCATTTCTTTGACCAAGTTGAGAAGCACTATGCAATTCG 564
Qy 517 CATGAGTACATGTTCCCTGATGCTGTGTAGGTTTCAGGCAATTTGTGAGGATGAAGGGTT 576
Db 565 AATTGAGTACATGTTTCCAGATGCTGTGTGAAGTTTCAAGCATTTAGTAAGGAACAAGGGTTT 624
Qy 577 ATTCTCTTTTACGAGATGGCAACCAAGAGTGTTCAGGAGTGAAGAGTGAAGGCTTTT 636
Db 625 ATTCTCTTTTACGAGATGGCAATCAGGAGTGTCTCGGTGTGAGGAAGGTGAGACCTTT 684
Qy 637 AAGGAGGCGCTTAAAGGTTCTCAGAGCATGATAAATGCTGAGAGGAAGAACCAAGTCACC 696
Db 685 GAGCGGGCTCTGAAGGGCTTCGGGCTGATCACTGGCCAAAGGAAGGATCAATCTCC 744
Qy 697 TGGTACTAGTCTGAATAACCGGTTGTTCAGGTTGTATCCGGCTTTTTCAGGGAATGGATGG 756
Db 745 AGGAGCAGGCTCTGAATTCAGTTGTTCAGGTGGACCCGGTTTTCAGGGATTTGATGG 804
Qy 757 TGGAAATGGAGCTTGTGTGAAGTGAACCTGTGTGCAAAATGTGAAGGCCATGACATATG 816
Db 805 TGGGCTTGAAGCTTATCAAGTGAATCCAAATGGCAATTTGTGAGGGAACAAGATGATG 864
Qy 817 GAACCTCTTTAGGACCATGAATGTGCTGTGAATTCCTGTCATGCAAAAGGATGTGTTTC 876
Db 865 GAAGTTCTTTCGAACATGATGATGTGCTGTGAATTCATTCATTCGATTCGAAAGGATACATCTC 924
Qy 877 CATTTGGTGTGAGCCCTGCACTAGGCTGTGTTTACCTGGGCAACATGAAGGAAGGAG 936
Db 925 TATTGGTGTGAGCTTTGCAAGGCCAGCTTCTTCTGCGCAACATGAGAGAGAAGGACG 984
Qy 937 GTGGTGTGGAGGATGCCAAGCTTAGGAATGTGCTTTCACAAAGGAATGTGAAGCA 996
Db 985 GTGGTGTGGAGGATGCCAAGCTTAGGAATGTGCTTTCACAAAGGAATGTGAAGCA 1044
Qy 997 GCAGAAAGAGGAGGATGTTAATGGAATGGGCTATCCCAATCCCATGCAAAATGGTGTATGC 1056
Db 1045 GGGTGTGAGACCCCACTTAATGGCAATGAAACGGGGCTGCCCATGCAAAATGG---GGC 1101
Qy 1057 TACACTGTCCCTGACATTTTCAACAGCCCGCAATGTATGTTTAACTTGAGCAGGACTGGAAT 1116
Db 1102 TGGCACTGTGCTGATATTTCAACTCCGAGAAATTTGGTCAATTTGAGCAGGCCCGCAAT 1161
Qy 1117 TGAGAAATTTGGCAAAATTTGAGGACCGAAGGAACCATGCTTGTGTGCTTTATGCAAC 1176
Db 1162 CGAGAACTTGTAAATTTAGAGAACCGAAGGAACCATGCTTGTGTGCTTCTACGCCCC 1221
Qy 1177 ATGGTGTCCCTACTGTCAGGCTATGAGGAATCTTATGTTGACTTATGACACACAAGTTAGC 1236
Db 1222 ATGGTGCAATTTTGCAAGGATGAGGACATCTTATGTTGAATTTGGCCATGAGTTAGC 1281
Qy 1237 AGGCTCAACAGGATGAAGTTTGAAATTTAGAGCAGATGAGGAAACAGAAAGAAATTCG 1296
Db 1282 AGGG---AGTGGAGTAAGGTGGGAAATTCAGGGCAGATGGCGATCAGAAAGGATTCCTC 1338
Qy 1297 AAAGATGTAAGTGAATTTGGAGCTTCCCTACGATATTTATTTTCCCAAGCAATTCGTC 1356
Db 1339 AAACAGGAATTAAGCTAGGAAGTTTTCACAAATCTTTCTTCCCTAAACACTCATC 1398
Qy 1357 TCGGCCAAACAATGAATGATCCCTCAGAAAAAGAGATGTTGATTTCTTGTATGCAATTTGT 1416
Db 1399 TCGCC---CAATCAGTACCCCTCTGAGAAAGGAGCTGGATTCATTTGATGACTTTTGT 1455
Qy 1417 AATGCTTAAAGTGA 1432
```

```
Db 1456 GAATGCTCTCCGGTAA 1471
|||||
RESULT 5
CRU63784 1753 bp mRNA linear PLN 15-FEB-2001
LOCUS Catharanthus roseus PAPS-reductase-like protein (par2neu) mRNA,
DEFINITION complete cds; nuclear gene for chloroplast product.
ACCESSION U63784
VERSION U63784.2 GI:12831473
KEYWORDS
SOURCE Catharanthus roseus (Madagascar periwinkle)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Gentianales; Apocynaceae; Rauvolfioideae;
Vincaceae; Catharanthus.
REFERENCE 1 (bases 1 to 1753)
AUTHORS Prior,A., Uhrig,J.F., Heins,L., Wiesmann,A., Lillig,C.H.,
Stoltze,C., Soll,J. and Schwenn,J.D.
TITLE Structural and kinetic properties of adenyllyl sulfate reductase
from Catharanthus roseus cell cultures
JOURNAL Biochim. Biophys. Acta 1430 (1), 25-38 (1999)
PUBMED 10082930
REFERENCE 2 (bases 1 to 1753)
AUTHORS Uhrig,J.F., Prior,A. and Schwenn,J.D.
TITLE Direct Submission
JOURNAL Submitted (03-FEB-2000) Biochemistry of Plants,
Ruhr-Universitaet-Bochum, Bochum 44780, Germany
REMARK Sequence update by submitter
COMMENT On Feb 15, 2001 this sequence version replaced gi:1488042.
FEATURES
source
1..1753
/organism="Catharanthus roseus"
/mol_type="mRNA"
/db_xref="taxon:4058"
1..1753
/genes="par2neu"
/evidence=experimental
106..1500
/genes="par2neu"
/function="reduction of adenyllyl sulfate (APS)"
/notes="C-terminus similar to pdi and beta-subunit of
4-hydroxylase, PDB Accession Number: 1MEK"
/codon_start=1
/evidence=experimental
/product="PAPS-reductase-like protein"
/protein_id="AAB05871.2"
/db_xref="GI:12831474"
translation="MALFTSTAIHGLSSSPQTKAAAFQSGFQPLDRPHTISPS
VNVSRRLAVKPINAEPKNESIVPSAATTVAPVEKVDVEDYKLADELQNASPLE
IMKSLAKGNDIAIAFSGAEDVALLEYAHLTRPFRVSLDTRLNPEYTKFPDITVE
KYQGHIEYMFDAVEQALVRSGKGLFSFYBDHGQCCRVKVRKVRRLRALUKRAWIT
GQRKQSPGTRSEIPVQVDVPFEGDGGVSLVKNVPANVEGKDIWNFLRAMDVPV
NTLHSGYVSIIGCEPTRPVLPGQHEREGRWMEADAKACEGLHKGNIKEETLNNGN
GAVNGSDPTIADIPTDNVTLSLRPGIENLLKBERREALVVLVAPWCRFCQAMEG
SVLEAEKLAGSGVKYKFKADGDQKAFQAQQLQLNSFTLILFFPKHSSKPIKYPSEK
RDVDSLMAFVNALR"
transit_peptide 106..282
variation
1027..1030
/genes="par2neu"
/notes="results in cysteine to tryptophan substitution"
/citations=[1]
/replaces="tgt"
1072..1075
variation
```

```
/gene="par2neu"  
/note="results in glutamine to aspartic acid substitution"  
/citation=[1]  
/replace="gac"  
1084..1086  
/gene="par2neu"  
/note="results in glutamic acid to glycine substitution"  
/citation=[1]  
/replace="gga"  
1093..1095  
/gene="par2neu"  
/note="results in asparagine to isoleucine substitution"  
/citation=[1]  
/replace="ata"  
1096..1098  
/gene="par2neu"  
/note="results in asparagine to isoleucine substitution"  
/citation=[1]  
/replace="ata"  
1099..1101  
/gene="par2neu"  
/note="frameshift results in glycine to tryptophan substitution"  
/citation=[1]  
/replace="tgg"  
1102..1104  
/gene="par2neu"  
/note="frameshift results in glycine to tryptophan substitution"  
/citation=[1]  
/replace="tgg"  
1105..1107  
/gene="par2neu"  
/note="results in asparagine to aspartic acid substitution"  
/citation=[1]  
/replace="gat"
```

## ORIGIN

```
Query Match 43.4%; Score 706.8; DB 15; Length 1753;  
Best Local Similarity 71.9%; Pred. No. 2.5e-175;  
Matches 1017; Conservative 0; Mismatches 367; Indels 30; Gaps 6;
```

```
QY 31 AATGGCCCTCGCTTCACTTCTCAATTC-----CGCACCACTCCACCTCCC 81  
DB 105 AATGGCTTGGCTTCACTTCTCAATTCGCAATTCATGGCTCTCTGCTTCTTCAATTGA 164  
QY 82 ATCATCGGAACCCAACTTCGCAAAATGGGTCAATTAGGATTTTCGAGAGGCCCATTTGG 141  
DB 165 ACAACAAAGACGACGCGCGAGTTTGGTTTCATTTTCAGCCTTTGGATCGGCCACATAC 224  
QY 142 AGGCGCGTTAATTTCAATTTATCTCAAGACGGAGCTTGTAAAGCCGTTTAAGCCGA 201  
DB 225 TATATCACTTTCGGTGAATGTTCTCGTAGACGTTTGGCCGTTAAGCCAATTAACGCTGA 284  
QY 202 ACCTCCAGCAAGATTCATTTGTTCTCTCGAGCAACACCATCGTGTGCTTCTGCTTC 261  
DB 285 GCCTAAAGGAATGAGTCAATTTGTCCTCGAGCAACACCGTTGCTCCGAGTTGA 344  
QY 262 TGAGA---CGAAAGAGGAAGATTTTGAACAGATAGCCAGTGTATCGCAATGCTTACC 318  
DB 345 AGAGAAAGTTGATGTGGAAGACTATGAGAAATTTGGCTGATGAGCTTCAAAATGCCCTCC 404  
QY 319 TCTTGAATCATGATAGAGCCCTCGACAAATTCGGGAACGACATAGCTATTGCTTCAG 378  
DB 405 TCTTGAGATCATGGAACAAGTCCCTGGCTAAGTTTGGCAATGATATAGCTATTGCTTCAG 464  
QY 379 TGGTCTCAAGATGTTGCTTTGATTGAGTATGCGAAATTCAGCGGTTCACCTTTAGGGT 438  
DB 465 TGGGGCCGAAGACGTTGCTTGTATAGAGTATGCACACTTGACTGCTGACCAATTCAGAGT 524  
QY 439 TTTCAGTTTGGACACTGGGAGACTGAACCCGAGAACTTATCAACTTTTGTGATCGGTTGA 498  
DB 525 ATTTAGCTTGGACACAGGAGATTGAATCCGAGACATATAAAATTTTGTGATACGTTGA 584
```

```
QY 499 GAAGCATTATGGAATTCGCATTTCAGTACATGTTCCCTGATGCTGTTGAGGTTTCAGGCATT 558  
DB 585 GAAACAGATATGGAATTCATATTCAGTACATGTTTCTGATGCTGTCGAAGTTTCAGGCTTT 644  
QY 559 GGTGAGGAGTAAAGGGGTTATTTCTTTCTACGAGATGGGCAACCAAGAGTGTTCAGAGGT 618  
DB 645 AGTTAGGAGCAAGGGGCTTTTCTTTCTACGAAGATGGCCACCAAGAGTGTCTCCGGGT 704  
QY 619 GAGAAAGGTGAGGCTTTAAGGAGGGCCCTTAAGGCTTCAGAGCATGGAATACTGCTCA 678  
DB 705 GAGGAAGGTGAGGCTTCGGAAGAGCCCTTAAGGGATTAAGGCTTCAGAGCATGGAATACTG 764  
QY 679 GAGGAAGAGCAGTCACTGCTACTAGTCTCGAAATACCGTGTTCAGGTTGATCCGGC 738  
DB 765 GAGGAAGATCAGTCTCTGCGCACTAGTCTGAAATTCAGTGTTCAGGTTGACCTGT 824  
QY 739 TTTTGAAGGAATGGAATTTGGAAGCTTGGTGAAGTGGAAACCTGTTGCAAAATGT 798  
DB 825 TTTTGAAGGAATGGAATGGAATGGAAGCTTGGTGAAGTGGAAATCTGTTGGCTTAATGT 884  
QY 799 GAAGGCGCATGACATATGGAACCTTCTTAGGACCATGAATGCTGCTGTGAATTCCTTGCA 858  
DB 885 GGAAGGAAGGATATCTGGAACCTTCTCGAGCCATGATGTCGGTGAACATTCATTCGA 944  
QY 859 TGCAAAAGGATATGTTTCCATTTGGTGTGAGCCCTGCACTAGGCTGTGTTTACCTGGGCA 918  
DB 945 CTCAAGAGTTATGTTCAATTTGGCTGTGAGCCGTGCAAGGCGCAGTCTCTTCAGGGCA 1004  
QY 919 ACATGAAAGGGAAGGGAGGTGGTGGGAGGATGCAAAAGCTTAAGGAATGTGCTTCA 978  
DB 1005 GCACGAGAGGGAAGGAGGTGGTGGGGAAGATGCAAAAGCAAAAGGAATGTGGTTACA 1064  
QY 979 CAAAGGAATGTAAGACGACGAAAGAGGAGGATGTTAAATGGAATGGGCTATCCCAATC 1038  
DB 1065 CAAAGGAATATCAAG-----GAGGAACCTCTGAATTAATATGGAATGGAAGTGT 1115  
QY 1039 CCATGCAAAATGGTGTGATGCTACCACTGTGCTGTGACATTTTCAACAGCCCGAATGTAGTTAA 1098  
DB 1116 TAATGGAATGGAAGTGAATGATCTATTG---CTGATATCTTTGATACCAATATGTAACGAG 1172  
QY 1099 CTTGAGCAGGACTGGAATTTGAGAAATTTGGCAAAATTTGGAGGACCGCAAAAGGAACCATGGCT 1158  
DB 1173 CTTGAGTAGGCTCGGATTTGAGAAATTTGCTGAAGTTGGAGGAAAGAGAGGAGCATGGCT 1232  
QY 1159 TGTGTGCTTTATCCACATGCTGCCCTACTGCTCCAGCTATGGAGGATCTTATGTTGA 1218  
DB 1233 AGTTGTACTTTTACGACCTTGGTGCCTGTTTTCGACGCAATGGAAGGATCATATCTTGA 1292  
QY 1219 CTTAGCAGCAAGTTAGCAGGCTCAACAGGCTCAAGGTTGGAAGTTTGAAGGATTTAGAGCAGATGG 1278  
DB 1293 ATTGGCAAAAGTTGGCTGGT---CTGGTGTAAAGTTTGAAGGATTTAAGGCAGACGG 1349  
QY 1279 AGAACAGAAAGATTTGCAAAAGTGAATGCAATTTGGGAAGCTTCCCTACGATATTAT 1338  
DB 1350 TGATCAGAAAGCAATTTGCACAGCAAGATTTGCAGCTGAATAGTCTTCCCTACAATTTCTTT 1409  
QY 1339 TTTCCCAAGCATTTGCTCGGCCCAACAATAAAGTATCCCTCAGAAAAAGAGAGATGTTGA 1398  
DB 1410 CTTCCCAAGCACTCATCTAAGC---CCATAAAATATCATCAGAGAAAGAGGATGTAGA 1466  
QY 1399 TTTCTTGATGCAATTTGTAATGCTTAAGATGA 1432  
DB 1467 CTCATTGATGCTTTTGTATGCTCGCATGA 1500
```

## RESULT 6

```
BT013254  
LOCUS Lycopersicon esculentum clone 134763R, mRNA linear PLN 11-MAY-2004  
DEFINITION Lycopersicon esculentum clone 134763R, mRNA sequence.  
ACCESSION BT013254  
VERSION BT013254.1 GI:47104669  
KEYWORDS FLI CDNA.  
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
```

```

ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 1835)
Kirkness, E.F., Wang, W. and Vazeille, A.
Direct Submission
Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
Medical Center Drive, Rockville, MD 20850, USA
FEATURES
Location/Qualifiers
source
1..1835
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
/clone="134763R"
/tissue type="callus tissue"
/note="TCADM70"

ORIGIN
Query Match 40.8%; Score 664.8; DB 15; Length 1835;
Best Local Similarity 70.4%; Pred. No. 3.1e-164;
Matches 974; Conservative 0; Mismatches 377; Indels 33; Gaps 5;

Qy 69 CTTCACCTTCCCATCATCGAACCACCACTTCGCAAAATTCGGATTCATTTAGGATTTGCG 128
Db |||||
Qy 142 CTTCACTCTTCTTATGAAACAACCCAAAGATATCCCAATTCGGGTACCTTTTCAGCCATTCG 201
Db |||||
Qy 129 AGAGGCCCATTCGGAGGCGCCGTTAATTTCAATTTATCTCAAGAGACGGAGCTTGTGTAAGC 188
Db |||||
Qy 202 ATAGGCTCAACTATTCGTCACACTGTTTGAATTTCTCGAGGCGGTTCGCGATGAGC 261
Qy 189 CCGTTAAACGCCGAACCTCCACGCAAGGATTCATTTGTTCTCTCGCAACCAACACCATCG 248
Db |||||
Qy 262 CATTTGATGCTGAAACCTAAGAGGAATGATTCATAGTTCCGTCAGCAGCTACCATCGTG 321
Qy 249 TTGCTTCTGCTTCTGAGAGGAAAGGNA---GATTTGAACAGATAGCCAGTCTCG 305
Db |||||
Qy 322 CTCCTGAGGTAGGAGAGAGTGTGAGGAGAGGACTTTTGAGAAATTCGCTAAGGAGCTC 381
Qy 306 ACAATGCTTCACTCTTGAATCATGATAGAGCCCTCGACAAATTCGGCAACACACATAG 365
Db |||||
Qy 382 AATAGCTTCCCTCTTGAGGTTATGCAACAGACATTCGAAATTTGGAGATGACATG 441
Qy 366 CTATTGCTTCACTGCTGCTGAGATGTTGCTTTGATGATGATGCGAAATTCAGCGGTC 425
Db |||||
Qy 442 CTATTGCTTTCAGTGGTGTGAAGATGTTGCTTTGATAGATGACGACATTTAACTGAC 501
Qy 426 GACCTTTAGGTTTTCAGTTTGGACACTGGGAGCTGGAACCCAGCAAACTTATCACTTT 485
Db |||||
Qy 502 GACCATACAGAGTATTCAGCCTTGATACCTGGAGGTTGAACCCCGAGACATACCAATTAT 561
Qy 486 TTGATCGGTTGAGAAGCATTTGCAATTCGCAATTCGATGATGATGCTGATGCTGTTG 545
Db |||||
Qy 562 TTGACAGCTGGAGAGCACTATGCAATTCGCAATTCGATGATGATGCTGATTCAGTTG 621
Qy 546 AGGTTACAGCAATTCGTTGAGGAGTAAAGGGTTATTTCTTTCTACGAGGATGGGACCAAG 605
Db |||||
Qy 622 AAGTTACAGGCTTGGTTAGGACCAAGAGGCTTTTCTTTCTTTCTATGAGGATGGCCACCAAG 681
Qy 606 AGTTTCAGGCTGAGAAAGTGAAGGCTTTAAGAGGCGCCCTTAAGGGTCTCAGAGCAT 665
Db |||||
Qy 682 AGTTTCAGGCTGTAAGGAAGGTTAGGCTTTTGAAGGAGGCTCTAAGAGGCTTACGCGCT 741
Qy 666 GGATAACTGTTGACAGAAAGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 725
Db |||||
Qy 742 GGATCAGGCGCCGTAAGATCAGTCCCTGGAATTCGATGATGATGATGATGATGATGAT 801
Qy 726 AGGTTGATCGGCTTTTGAAGGAAATGATGATGATGATGATGATGATGATGATGATGATG 785
Db |||||
Qy 802 AGGTTGATCGGCTTTTGAAGGAAATGATGATGATGATGATGATGATGATGATGATGATG 861
Qy 786 CTGTTGCAATGTGAAGGCGCATGACATATGGAATCTTCTTAGGACCATGATGCTGCTG 845
Db |||||

```

```

Db 862 CTGTGGCTAATGTGGACGGAAAGATATTTGGAACTTCTCTGCTGCCATGAATGTGCTG 921
Qy 846 TGAATTCCTTCATGCAAAAGGATATGTTTCATTTGGGTGTGAGCCCTGCACTAGGCTG 905
Db |||||
Qy 922 TGAATTCATTCATTCACAGGATATGATTCATTTGATGCAACCTTGCACAGGCCAG 981
Qy 906 TTTTACTGCGCAACATGAAGGGAAGGAGGTGGTGGTGGAGGATGCCAAAGCTAAGG 965
Db |||||
Qy 982 TTCTACCAAGGCAACACGAGAGAGGGAAGATGTTGGTGGGAAGATGCCAAGGCCAAG 1041
Qy 966 AATGCTGCTTCACAAAGGAATCTAAAGCAGCAGAAAGAGAGGAGGATGTTTATGGAATG 1025
Db |||||
Qy 1042 AGTGTGGCTTGCAACAAGGGCAACATCA-----AGGATGAATCTGTAAATG 1086
Qy 1026 GGCTATCCCAATCCATGCAAAATGGTGTATGCTACACCTGTGCTGCTGATATTTTATGATCA 1085
Db |||||
Qy 1087 GCGCTGCCCAACAAA-----TGGTACTGCTACCGTTGCTGATATTTTATGATCA 1137
Qy 1086 CGAATGTAGTTAATTCAGCAGGAGTGGAAATTTGAGAAATTTGGCAAAATTTGAGAGCA 1145
Db |||||
Qy 1138 AGGACATTTGTACCTTCAGTAAAGCTGGAGTTGAGAACCTAGTAAATTTGGAAGACCGAA 1197
Qy 1146 AGGAACCATGCTGTTGCTTTATGCAACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1205
Db |||||
Qy 1198 GAGAGCCTTGGCTGCTGTTGCTTTTATGACCTTGGTCCCAATTTTGGCAGCAATGGAAG 1257
Qy 1206 AATCTTATGTGACTTAGCAGACAAAGTTAGCAGGCTCAACAGGATGAAGGTTGGAAAT 1265
Db |||||
Qy 1258 GATCTATGTTGAATTCGCTGAGAAATTTGGCTGCTGTT---CTGGTGTGAAAGTAGGAAAT 1314
Qy 1266 TTAGAGCAGATGGAGAAACAGAAAGAAATTTGCAAGAGTGAACCTCAATTTGGAAAGCTTCC 1325
Db |||||
Qy 1315 TCAGGCGAGATGGTGGACAGAAAGCAATTTGCAAGAAAGAAATTCAGCTTGGCAGCTTCC 1374
Qy 1326 CTAGCATATTTATTTTCCCAAGCACTTCGCTCGGCCCAACAAATTAAGATATCCCTCAGAAA 1385
Db |||||
Qy 1375 CTACAATACTCTTCTTCCCAAGCACTTTCAAAG---GCCATTAAGTACCTTTTCAGAGA 1431
Qy 1386 AGAGAGATGTTGATTCCTTCATGAGCAATTTGTAATGCTTAAAGATGAGGATATCAGAAA 1445
Db |||||
Qy 1432 AGAGGAGCTAGACTCTTCTGCTGCTTGTGGAATGCTCTCAGATGAAGAAAGCAATGCTG 1491
Qy 1446 TTTT 1449
Db |||||
Qy 1492 TTTT 1495

RESULT 7
AY568717
LOCUS
DEFINITION Lycopersicon esculentum adenyl-1-sulfate reductase mRNA, complete
cds.
ACCESSION AY568717
VERSION AY568717.1 GI:51457939
KEYWORDS
SOURCE
ORGANISM Lycopersicon esculentum (Solanum lycopersicum)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 1386)
Coker, J.S., Vian, A. and Davies, E.
Identification, accumulation, and functional prediction of novel
tomato transcripts systemically upregulated after fire damage
Physiol. Plantarum 124 (3), 311-322 (2005)
2 (bases 1 to 1386)
Coker, J.S., Vian, A. and Davies, E.
Direct Submission
Submitted (08-MAR-2004) Botany, North Carolina State University,
Gardner Hall, Campus Box 7612, Raleigh, NC 27695, USA
FEATURES
Location/Qualifiers
source
1..1386
/organism="Lycopersicon esculentum"

```

```

/mol_type="mRNA"
/db_xref="taxon:4081"
1..1386
CDS
/codon_start=1
/product="adenyl-lyl-sulfate reductase"
/protein_id="AAU03359.1"
/db_xref="GI:51457940"
/translation="MAUTFTSSSAIHGSLSSSSSYRQPKVQLGTQPLDRPOLLS
TVLNSRRSAVKPLIAPRKNDSIVPSAATIVAPGVESVEAEDEFKALQNASPL
EVMKALKEFGDDITAFSGAEDVALIEYAHLTGRPYRFSLDTRLNLPETYQLFDTV
EKHYGIRIEMFDPVSVOALVRKGLFSEFDGHQBCRRVKRVPRLRALKGLRWVI
TGRKQDSPTRESPIPOVDPSFEGLDGGAGSLVKNPVANVDKDIWNFLRANVP
VNSHSQTVSIGEPECTRPVLPQCHERERHWEADAKAKECGIHKNGIKDETWNGA
QTNGTATVADIFDKIIVITLSKPGVENLVKLEDRRPLWVLYAPWQCFQAMGSGTV
ELAEKLAGSGVKGFRADGQKAFAPAEELQLGFPTILFFPKHSSKAIKYPSKRDV
DSLAFVNALR"
transit_peptide 1..213
/note="chloroplast transit peptide"
/evidence=not_experimental
misc_feature
346..963
/note="Region: phosphoadenosine phosphosulfate reductase
domain"
misc_feature
1048..1377
/note="Region: thioredoxin domain 2"
ORIGIN
Query Match 40.7%; Score 663.8; DB 15; Length 1386;
Best Local Similarity 70.7%; Pred. No. 5.7e-164;
Matches 967; Conservative 0; Mismatches 367; Indels 33; Gaps 5;
QY 69 CTTCCACCTTCCCATCATCGGAACCAACCTCCCAATTTGGGTCAATTAGGATTTCGG 128
DB 50 CTTTCATCTTCTTCTATGAAACCAACCAAGATATCCCAATTCGGTACCTTTTCAGCAATGG 109
QY 129 AGAGGCCATTTGGAGCGCGTGAATTTCAATTTATCTCAAGACGGAGCTTGTAAAGC 188
DB 110 ATAGCCCTCAACTATTGTGCTCACTGTTTGAATCTCGAGGGGTTCGGCAGTGAAGC 169
QY 189 CCGTTAAACCGCGAACCCTCCACGAGGATTCATTTGTTTCTCTCCAGCAACCAACCATCG 248
DB 170 CATTGTATGCTGAACCTAAGAGGAATGATCAATAGTTCCGTCAGCAGCTACCACTCGTGG 229
QY 249 TTGCTTCTGCTTCTGAGAGGAAGAGGAA---GATTTGACAGATACCCAGTGATCTCG 305
DB 230 CTCCTGAGGTAGGAGAGGATGTTGAGGAGAGGACTTTGAGAAATTTGGTAAAGGAGCTTC 289
QY 306 ACAATGCTTCACTCTTGAATCATGGATAGAGCCCTCGACAAATTCGCGCAACGACATAG 365
DB 290 AAAATGCTTCCCTCTTGAGGTATTTGACAAAGCACTTGAGNAATTTGGAGATGACATTG 349
QY 366 CTATTGCTTTCAGTGGTGTGAAGATGTTGCTTTGATTTGAGTATCGGAAATTTGACGGGTC 425
DB 350 CTAATTGCTTTCAGTGGTGTGAAGATGTTGCTTTGATAGTATGACCAATTTAATCGGAC 409
QY 426 GACCTTTAGGGTTTTCAGTTTGGACACTGGAGACTGAACCCAGAACTATCAACTTT 485
DB 410 GACCATACAGAGTATTCAGCTTGTATCTGGGAGGTGTAACCCGAGACATACCAATTAT 469
QY 486 TTGATCGGTTTCAGAGCAATATGAATTCGCAATGAGTACATGTTCCCTGATGCTGTG 545
DB 470 TTGACACAGTGGAGAGCACTATGCAATTCGCAATGATATGATTTCCCTGATTCAGTTG 529
QY 546 AGGTTTCAGGCAATTTGGAGGATTAAGGGGTTATTTCTTTTCACGAGGATGGCCACCAAG 605
DB 530 AAGTTTCAGGCGTTGTTAGGACCAAGGGCTTTCTCTTCTATGAGGATGGCCACCAAG 589
QY 606 AGTGTTCAGGCTGAGAAAGGTGAGCCCTTAAAGGAGCCCTTAAGGCTCTCAGAGCAT 665
DB 590 AGTGTTCGCGTTAAGAGAGGTGAGCCCTTTGAGGAGAGCTCTAAGGGCTTACGCGCT 649
QY 666 GGAATACGCTCAGAGGAAGACCACTGACCTGGTACTAGGTCTGAAATACCGGTTGTTTC 725
DB 650 GGAATACAGGCCGCTAAGATCAGTCCCTTGGAACTCGATCAGAAATCCCCAATTTGTTTC 709

```

```

QY 726 AGTTGATCGGCTTTTGAGGGAATGATGTTGGAATTTGGAAGCTTGGTGAAGTGGAAACC 785
DB 710 AGTGGACCTTCTTTTGGGGTGGATGGCGGTGCTGGTAGCTTGGTGAAGTGGAAACC 769
QY 786 CTGTTGCAATTTGGAAGGCGCATGACATATGGAATCTTCTTTAGGACCATGAATGTGCTG 845
DB 770 CTGTGGCTAATGTGGACGGAAGATATTTGGAATCTTCTTGGTGGCATGAATGTGCTG 829
QY 846 TGAATTCCTTGCATGCAAAAGGATATGTTTCATTTGGTGTGAGCCCTGCACTAGGCTG 905
DB 830 TGAATCTATTGCAATTCACAGGATATGATTCATTTGGATGCGAACTTTGCAACAAGGCCAG 889
QY 906 TTTTACCTGGGCAACATGAAGGGAAGGAGTGTGTGGGAGATGCCAAAAGCTAAGG 965
DB 890 TTTACACAGGCAACACAGAGAGAGGGAAGATGTTGTGGGAAGATGCCAAGGCCAAG 949
QY 966 AATGTGTGTTTCAAAAAGGAAATGTAAGACGACGAGAAGAGGAGATGTTAATGGAATG 1025
DB 950 AGTGTGGCTTGCACAGGGCAACATCA-----AGGATGAATCTGTAATG 994
QY 1026 GGCTATCCCAATCCCATGCAAAATGGTGAATGCTACCATGCTGCTGACATTTTCAACAGCC 1085
DB 995 GCGCTGCCCAACAAA-----TGTACTGTCTACCGTTGCTGATATTTTGTATCCA 1045
QY 1086 CGAATGTAGTTAACTTGGACGAGGACTGGAATTTGCAAAATTTGCAAAATTTGAGGAGCCGAA 1145
DB 1046 AGGACATTTGTACCTTGGATGAAGCTGGATGAGACCTAGTAAATTTGAGAGCCGAA 1105
QY 1146 AGGAACCATGCTTGTGCTTTATGCACCATGTTGCTGCTTCTGCTGAGGCTATGGAGG 1205
DB 1106 GAGAGCTTGGCTGTTGTTCTTTATGCACCTTGGTGCCTTTTGGCAGCAATGGAAG 1165
QY 1206 AATCTTATGTTGCTTAGCAGACAAAGTTAGCAGGTTCAACAGGATGAAGTTGGAAT 1265
DB 1166 GATCTATGTTGAATTTGGCTGAGAAAGTTGGCTGTT---CTGGTGTGAAGTAGGGAAT 1222
QY 1266 TTAGACGAGTGGAGACAGAAAGAAATTTGCAAGAGTGAATGCAATTTGGAAGCTTCC 1325
DB 1223 TCAGGGCAGATGGTGACCAAGAACATTTGACAGAAAGATTTGCGAGCTTGGCAGCTTCC 1282
QY 1326 CTACGATATTTATTTTCCCAAGCAATTCGTCTCGGCCAACATAAAGATATCCCTCAGAAA 1385
DB 1283 CTACATATCTTCTTCTCCCAAGCACTCTTCAAAG---GCCATTAACTACCTTCAGAGA 1339
QY 1386 AGAGAGATGTTGATTCCTTGTATGGCAATTTGTAATGCTTAAAGTGAATGA 1432
DB 1340 AGAGGAGCTAGACTCTCTTGTGCTTGTGATGCTTGTGAATGCTCTCAGATGA 1386

RESULT 8
SCJ001208
LOCUS SCJ001208 1608 bp mRNA linear PLN 21-MAY-1999
DEFINITION Brassica juncea mRNA for APS reductase.
ACCESSION AJ001208
VERSION AJ001208.1 GI:4033576
KEYWORDS APS reductase.
SOURCE Brassica juncea
ORGANISM Brassica juncea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1
Heiss, S., Schafer, H.J., Haag-Kerwer, A. and Raesch, T.
Cloning sulfur assimilation genes of Brassica juncea L.: cadmium
differentially affects the expression of a putative low-affinity
sulfate transporter and isoforms of ATP sulfurylase and APS
reductase
Plant Mol. Biol. 39 (4), 847-857 (1999)
10350097
REFERENCE 2 (bases 1 to 1608)
AUTHORS Heiss, S.
TITLE Direct Submission
JOURNAL
PUBMED
REFERENCE 2 (bases 1 to 1608)
AUTHORS Heiss, S.
TITLE Direct Submission

```

JOURNAL Submitted (15-JAN-1998) Heiss S., Ruprecht-Karls-Universitaet  
Heidelberg, Botanical Institute, Im Neuenheimer Feld 360, D-69120,  
GERMANY

## FEATURES

Source Location/Qualifiers  
1..1608  
/organism="Brassica juncea"  
/molecule="mRNA"  
/cultivar="Vitasso"  
/db\_xref="taxon:3707"  
/tissue\_type="root"  
1..1608  
/gene="apsr8"  
30..1424  
/gene="apsr8"  
/note="putative mitochondrial or plastidic transit  
peptide"  
/codon\_start=1  
/product="APS reductase"  
/protein\_id="CAA04611.1"  
/db\_xref="GI:4033577"  
/db\_xref="GOA:Q92P22"  
/db\_xref="UniProt/TreMBL:Q92P22"  
/translation="MALAIVSSSSSAISTSSPPSELKVSAPRIGSLRLSDRVNVS  
SVLSLGGKSSVKPLNQVIAKESVGSVASEVTERKQDVVEDEFEELAKSLTASP  
LEIMDKALETFGNIDIAFAGAEVALIEYAHLTGRFVPSLDITPELRYELFDT  
VEKHGTRIRIEMFPDAVEQVALVRNKLGFSEYEDHQCCEIRKVRPLRLALKGLRAW  
ITGQRKQSPGTSSEIPVQVDPVPEGLDGGSLVKVNPVAVNGNDVNFRLTMDV  
PVNTLHAAGVSVIGCEPCTRAVLPGOHEREGRWEDAKAKEGCLHKGNIKSSNGNA  
TANVNGTSVVDIPEKSNVSVLSROGIENLMLENKEAWIVLYAPPCFPQMBEA  
SPBELADLGGSGVKVAKFRADGQKEFAKRELQLGSPFILLVFPKNSRPIKTPSEK  
RDVSLTSFLNLVR"

transit\_peptide 30..194  
/gene="apsr8"  
/note="putative mitochondrial or plastidic transit  
peptide"

mat\_peptide 195..1421

/gene="apsr8"  
/product="APS reductase"  
/note="putative mitochondrial or plastidic transit  
peptide"

## ORIGIN

Query Match 40.7%; Score 663.6; DB 15; Length 1608;  
Best Local Similarity 68.7%; Pred. No. 6.4e-164;  
Matches 985; Conservative 0; Mismatches 419; Indels 30; Gaps 4;

QY 32 ATGCGCCCTCGCTTCACTTCTTCAATTTCCGACCAACTTCCACCTTCCCATCATCGAA 91  
DB 42 ATCAACGTTTCTTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 101  
QY 92 CCCAAACT-----TCCGCAAAATGGGTCAATTAGGATTTCCGAGAGGCCATTGGAGGC 145  
DB 102 CTCAAGTTTCAGCTCCACGGATCGGTTCGTGGGTTATCGGATCGTGTCAATGTCTCA 161  
QY 146 GCCGTTAATTCAATTTATCTCAAGACGGAGCTTGGTAAAGCCCGTTAAACGCCGACCT 205  
DB 162 TCGGTGCTCTGAGTCTATCGGGAACGATCTCTCGGTGAAGCTCTGAATGTGCACTCA 221  
QY 206 CCAGCGAAGATTCATTGTTCTCTCGCAACCAACCACTGTTGCTTCTGCTTCTGAG 265  
DB 222 ATTGCAAGAGAGTGGTGGTTCATGTTGTTCTATGTTGCTTCTGAGTTACAGAAACAAGATGT 281  
QY 266 ACAGAAAGAGGAAGATTGTAACAGATAGCCAGTGTATCTGCACAAATGCTTCCACCTCTTGAA 325  
DB 282 GTGGAGTGAAGACTTTGAGGAGCTAGCAAGAGTCTAGAGACTGCTTCTCTCTTCTTCTGAG 341  
QY 326 ATCATGATAGAGCCCTCGCAAAATTCGCAACGACATAGCTATTGCTTTTCAGTGGTGTCT 385  
DB 342 ATTATGACAAAGCTCTTGAGACATTCGGAACGACATCGCAATCGCTTTTAGTGGAGCA 401  
QY 386 GAAGATGTTCTTGTAGTATGCAAAATTCGCAATTCAGCGGTTCAGCCCTTTAGGGTTTCAGT 445  
DB 402 GAAGACGTTGCTCTCATCGAGTATGCTCATCTTACTGGAAGACCCCTTCAGGGTTTCAGT 461

RESULT 9  
AP004494/c  
LOCUS

AP004494

103129 bp

DNA

linear

PLN 22-JUL-2003





Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A., and Ecker, J.R.  
 Arabidopsis ORF clones  
 Unpublished  
 2 (bases 1 to 1377)  
 Kim, C.J., Chen, H., Cheuk, R., Meyers, M.C., Shinn, P., Ban, J., Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayaishizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A., and Ecker, J.R.  
 Direct Submission  
 Submitted (11-MAY-2002) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA  
 RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA': Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayaishizaki, Y., and Shinozaki, K.)  
 The Salk, Stanford, PGC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Kim, C.J., Chen, H., Cheuk, R., Meyers, M.C., Shinn, P., Ban, J., Bowser, L., Chang, E., Dale, J.M., Goldsmith, A.D., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.  
 Kim, C.J. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.  
 Location/Qualifiers  
 1. .1377  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="4"  
 /clone="U13952"  
 /ecotype="Columbia"  
 /note="This clone is in pUNI 51"  
 1. .1377  
 /note="PRH26 protein"  
 /codon\_start=1  
 /product="AT4g21990/FLN20\_90"  
 /protein\_id="AAW65364.1"  
 /db\_xref="GI:21655305"  
 /translations="MALAINVSSSSAITSPPSSDLKVTIKIGSLRLNRTNVSAA  
 SLSLGRSSVKALNVQITKESIVASEVTEKLDVVEDEPELAKRLNLSAPLEIMD  
 KALEKGNDAIAFSAEDVALIEYHLGRPRVSLDTGRNLPETRIFDIVFKHY  
 GRIEYFNPDAVQALVRNKGILFSYEDHQECIRKVRPLRLAKGLRAVITQOR  
 KDQPSGRSIPVQVDPVEGDGVLGKWPVNVENDNVNLFRLMTDPVNTL  
 HAAGYVIGCEPCTVNLPSQHEREGRWMEADAKCEGLHKNIKENTNGNATNVN  
 GTASVADIENSENVNVLROGIEMLKLENKEATVILVYAPWCPQCAEASFDELA  
 DKLGSGVKVAKFRADGQDKFAKELQLGSPFTILVFPKNSRPRIKPSKRDVSL  
 TSFLNLR"  
 Query Match 40.3%; Score 657.2; DB 15; Length 1377;  
 Best Local Similarity 69.5%; Pred. No. 3.2e-162;  
 Matches 966; Conservative 0; Mismatches 388; Indels 36; Gaps 4;  
 31 AATGGCCCTCGCTTTCATTCTTCAATTTCCGACCAACTTCCACTTCCCATCATCGGA 90  
 |||||  
 12 AATCAACGTTTCTTCT 71  
 |||||  
 91 ACCCAAACTCCCAAAATGGGTCAATTAGGATTCGGAGAGCCCAATTGGAGGCCCGT 150  
 |||||  
 72 TCTCAAAGTAACAAAATCGGATCATTTGAGGTTATTGAATCGTACCAATTTCTCTCGGC 131  
 |||||

QY 151 TAATTTCAATTTATCTCAAGAGCGGAGCTTGGTAAAGCCCGTTAAACGCCGAACCTCCACG 210  
 |||||  
 DB 132 TTCTCTGAGTTTGTCCGGGAAGAGATCTCTCCGTGAAGCTCTTAATGTGAATCAATTAC 191  
 |||||  
 QY 211 CAAGGATTCATTTGTTCTCTCCGAGAGCAACACCATCGTTGCTTCTGCTCTGAGAGAA 270  
 |||||  
 DB 192 AAAGGATTCATTTGTTCTCTGAGGTTACAGAGAAGCTAG-----ATGTGGTGA 242  
 |||||  
 QY 271 AGAGGAAGATTTTGAACAGATAGCAGTATCTCGCAATGCTTCCACTCTCTGGAATCAT 330  
 |||||  
 DB 243 AGTTGAAGATTTTGAAGAACTAGCAAGAGATTAGAGAATGCTTCTCTCTTTGAGATCAT 302  
 |||||  
 QY 331 GGATAGAGCCCTCGACAAATTCGGCAACGACATAGCTATTGTCCTTCACTAGTGTGCTGAAGA 390  
 |||||  
 DB 303 GGATAAGCTCTTGAAGAGTTTGGGAATGACATGCTATTGTCCTTTAGTGGAGCTGAAGA 362  
 |||||  
 QY 391 TGTGCTTGAATTCGAGATTCGGAATTTGACGGGTGACCCCTTTAGGGTTTTTCAGTTTGA 450  
 |||||  
 DB 363 TGTGCTCTGATTTGAGTATGCTCATTTAACTGGAAGACCTTATAGGTTATTCAGTTTGA 422  
 |||||  
 QY 451 CACTGGGAGACTGAACCCAGAACTTATCAACTTTTGTATGCGGTTGAGAGGATTTATGG 510  
 |||||  
 DB 423 TACAGGAGATTGAATCCAGAAACATATAGACTCTTCGATACCGTGGAGAAACATTACGG 482  
 |||||  
 QY 511 AATTGCGATTGAGTACATGTTCCCTGATGCTGTTGAGGTTTCAGGCATTTGTTGAGAGTAA 570  
 |||||  
 DB 483 TATTCGAGTTGAGTATATGTTCTCTGATGCTGTTGAGGTTTCAAGCTTTAGTTAGGAACA 542  
 |||||  
 QY 571 GGGGTTATTTCTTTTACAGAGATGGGCAACGAAGATGTTGAGGGTGAAGAGGTGAG 630  
 |||||  
 DB 543 GGGTTGTTCTCTTCTATGAAGATGGTCAACAGGAGTGTTCGCTATTATGAAGAGGTGAG 602  
 |||||  
 QY 631 GCCTTTAAGGAGGCGCTTAAGGCTCTCAGAGACATGATTAACCTGTCAGAGGAAGACCA 690  
 |||||  
 DB 603 ACCTTTGAGCGGTGCGTTGAAGGTTTACGTGCTTGGATCCTGCTCAGAGGAAGATCA 662  
 |||||  
 QY 691 GTCACCTGTGTAGGTCTGAAATACCGGTTGTTTCAGGTTTGATCCGGCTTTTGAAGGGAAT 750  
 |||||  
 DB 663 ATCACCGGAGACAAGATCGGAGATTCCTGTTGTTTCAGGTTGATCCAGTGTTTGAGAGATT 722  
 |||||  
 QY 751 GGATGTGGAATTTGGAAGCTTGTGTAAGTGAACCCCTGTTTGAAGTGAAGGCCCATGA 810  
 |||||  
 DB 723 AGATGTGGAGTTGGTAGTTGCTGTAAGTGAATCCGTTGCTTAATGTTCAAGGGAATGA 782  
 |||||  
 QY 811 CATATGGAATCTCTTAGGACCATGATGTCCTGTGAATTCCTTGCATCAAAAGGATA 870  
 |||||  
 DB 783 TGTTTGAATTTCTTGAGGACGATGATGTTCTCTGCAACACACGCTTCACGCTGCGGGTTA 842  
 |||||  
 QY 871 TGTTTCCATTGGGTGTGAGCCCTGCACCTAGGCCCTGTTTACCTGGGCAACATGAAGGGA 930  
 |||||  
 DB 843 TGTTTCAATCGGATGTGAGCCATGCACGAGAGCGGTTTTCGCGGTCAGCATGAGAGAGA 902  
 |||||  
 QY 931 AGGAGGTGTGTGGAGAGATGCCAAAGTAAGGAATATGGTCTTTCACAAAGGAATGT 990  
 |||||  
 DB 903 AGGAGATGTGTGGGAAAGATGCTAAGGCTAAAGAGTGTGGGCTTTCACAAAGGGAATAT 962  
 |||||  
 QY 991 AAGACGACAGAAAGAGAGGAGATGTTAATGGAATGGGCTATCCCAATCCCATGCAATGG 1050  
 |||||  
 DB 963 CAAGGAGAATACCAATGGAAACCAACCGCTAATG-----TCAA 1001  
 |||||  
 QY 1051 TGATGCTACCACTGTGCTGACATTTTCAACAGCCCGGAATGTAGTTAACTTTGAGCAGGAC 1110  
 |||||  
 DB 1002 CGGACAGCCTCAGTCCGCGATATCTTCAATACGAGAGATGTTGTGNACTTTGAGCAGGCA 1061  
 |||||  
 QY 1111 TGAATTTGAATTTGGCAAAATTTGAGGACCGAAAGAACCATGGCTTGTGTGCTTTA 1170  
 |||||  
 DB 1062 AGGGAATTGGAATTTGATGAAGTTAGAGAACCGGAAAGAGGCTTGGATCGTGTGCTTTA 1121  
 |||||  
 QY 1171 TGCACCATGGTCCCTACTGCGAGGCTATGGAGGATCTTATGTTGACTTTAGCAGACAA 1230  
 |||||  
 DB 1122 TGCACCGGTGTGCGCGCTTTTGTTCAGGCAATGGAAGCTTCATTCGATGAATTTGGCGGATNA 1181  
 |||||

```

Qy 1231 GTTAGCGGTCACAGGATGAAGTTGGAAATTTAGACGATGAGAGCAAGA 1290
Db 1182 GCTGGGAGG---AAGTGGTGTGAAGGTTGTAAGTTTAGAGCTGATGGTACCAAGAGGA 1238
Qy 1291 ATTGCAAGAGTGAACTGCAATTTGGGAAGCTTCCCTACCATATATTTTCCCAAGCA 1350
Db 1239 CTTTGCCAAAGAGGTTGAGCTTGGAGCTTCCCAACATATCTTGTGTTCCCAAGAA 1298
Qy 1351 TTGCTCTCGGCCAACATAAGTATCCCTCAGAAAAGAGAGATTTGATTCCTTGATGGC 1410
Db 1299 CTCCTTCAAG---ACCAATCAAGTATCCATCAGAGAGAGAGATGTTGATTCCTTGACATC 1355
Qy 1411 ATTGTGAAT 1420
Db 1356 TTTCTTGAAT 1365

```

```

RESULT 11
AY093319
LOCUS AY093319
DEFINITION Arabidopsis thaliana PRH26 protein (At4g21990) mRNA, complete cds.
ACCESSION AY093319
VERSION AY093319.1 GI:20259941
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopses.
1 (bases 1 to 1555)
Southwick,A., Karlin-Neumann,G., Nguyen,M., Lam,B., Miranda,M.,
Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,
Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,
Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
Submitted (26-MAR-2002) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
e-mail for correspondence: arab@sequence.stanford.edu

```

```

RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Southwick,A.,
Nguyen,M., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J.,
Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K.,
Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P.,
Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Southwick,A., (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed
equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.
(SSP/Stanford) contributed equally to this work as PIs.

```

```

FEATURES
    source
        1..1555
        /organism="Arabidopsis thaliana"
        /mol_type="mRNA"
        /db_xref="taxon:3702"
        /chromosome="4"
        /clone="U16653"
        /ecotype="Columbia"
        /note="This clone is in pENTR/SD-dTopo This is a cloned
        PCR product using RIKEN clone RAFL08-16-K20 (AY062665) as
        a template"
    gene
        1..1555
        /gene="At4g21990"
        /note="synonym: FIN20.90"
    CDS
        1..1377

```

## ORIGIN

```

Query Match 40.3%; Score 657.2; DB 15; Length 1555;
Best Local Similarity 69.5%; Pred. No. 3.2e-162;
Matches 966; Conservative 0; Mismatches 388; Indels 36; Gaps 4;

Qy 31 AATGGCCCTCGCTTCACTTCTTCAATTTCCGACCAACTTCCACCTTCCCATCATCGGA 90
Db 12 AATCAACGTTTCTTCAATCTTCTTCTGCGATCTCAAGCTCTAGCTTCCCTTCTTCA 71
Qy 91 ACCCAAACTTCCGCAAAATGGGTCAATTAGGATTTCCGAGAGGCCCATTTGAGGCGCGT 150
Db 72 TCTCAAGTAAACAAAATCGATCATTTAGGTTATTGAATCGTACCAATGTCCTGCGGC 131
Qy 151 TAAATTTCAATTTATCTCAAGACGAGCTTGGTAAAGCCGTTAAACCGGAACTCCACG 210
Db 132 TTCTCTGAGTTTGTCCGGAAGAGATCTCTCGTGAAGCTCTTAATGTGCAATCAATPAC 191
Qy 211 CAAGGATTCATTTGTTCTCTCGCAGCAACAACCATCGTTGCTTCTGCTGAGACGAA 270
Db 192 AAAGGAATCCATTTGTTCTCTGAGGTTACAGAGAGCTAG-----ATGTGGTGA 242
Qy 271 AGAGGAAGATTTTGAACAGATAGCCAGTGTCTCGACAATGCTTCACTTCACTTTGAAATCAT 330
Db 243 AGTTGAAGACTTTGAGGAACATGACAAGAGATTAGAGATGCTTCTCTCTTGAGATCAT 302
Qy 331 GGATAGAGCCCTCGACAAAATTCGCAACGACATAGCTATTGCCCTTCACTGTTGTTGAAGA 390
Db 303 GGATAAAGCTCTTCGAGAAGTTTGGGAATGACATGCTATTGCTTTAGTGGAGCTGAAGA 362
Qy 391 TGTGCTTTGATTGATGAGTAAATTTGACGGGTCGACCTTTAGGGTTTTCAGTTTTCGA 450
Db 363 TGTGCTCTGATTGATGATGCTCAATTTAACTGGGAAGCTTTATAGGGTATTTCAGTTTGA 422
Qy 451 CACTGGGAGACTGAACCCAGAAAACCTTATCAACTTTTGTATGCGGTTGAGAAGCATTTATGG 510
Db 423 TACAGGGAGATTGAATCCAGAAAACATATAGACTCTTCGATACCGTGGAGAAACATTAAG 482
Qy 511 AATTGCGCATTCAGTACATATGTTCCCTGATGCTGTTGAGGTTTCAAGGATTTGGTGAGGAT 570
Db 483 TATTGCGATTGAGTATATGTTTCTGATGCTGTTGAGGTTTCAAGCTTTAGTTAGGAACAA 542
Qy 571 GGGGTATTCTTCTTCTACGAGGATGGCAACAGAGGTTTGCAGGGTTGAGAAAGTGAAG 630
Db 543 GGGTTGTGTTCTTTCTATGAAGATGGTCAACAGAGGTTGTCCTGCTATTAGAAAGGTGAG 602
Qy 631 GCCTTTAAGGAGGGCCCTTAAAGGGTCTCAGAGCATGGATACTGCTCAGAGGAAAGACCA 690
Db 603 ACCTTTGAAGGCTGGGTTGAGGGTTTACGTGCTTGGATCACTGCTCAGAGGAAAGATCA 662
Qy 691 GTCACCTCGTACTAGGTTCTGAATAACCGGTTGTTCAAGTTGATTCGGGCTTTTGGAGGAAT 750
Db 663 ATCACCCGGGCAAGATCGGAGATTCCTGTTGTTTCAAGGTTGATTCAGGTGTTTGAAGGAT 722
Qy 751 GGATGTTGGAATTTGAAGCTTTGGTGAAGTGAACCTCTGTTGCAAAATGTTGAAGGCGCATGA 810
Db 723 AGATGTTGGAGTTGGTAGTTGTTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 782
Qy 811 CATATGGAACTTCTCTTAGGACCATGAATGTGCTGTGAATTCCTTTGTCATGCAAAAGGATA 870

```

Db 783 TGTTCGAAATTTCTTGAGGACGATGGATGTTCTGTCAACACGCTTCACGCTCGCGGTTA 842  
 Qy 871 TGTTCATTTGGGTGAGCGCTCCTCAGCTAGAGCCCTGTTTACCTCGGCAACATGAAAGGGA 930  
 Db 843 TGTTCATCGGATGTCAGCCATGTCAGAGAGCGGTTTTCGCGGTGAGATGAGAGGA 902  
 Qy 931 AGGAGTGTGTGGGAGAGTGCACAAAGCTAAGGAATGTGCTTTCACAAAGGAATGT 990  
 Db 903 AGGAGATGTGTGGGAGAGTGTAAAGGCTTAAAGAGTGTGGCCCTTCACAAAGGAATAT 962  
 Qy 991 AAAGCAGCAGAAAGAGAGAGATGTTAATGGAATGGGCTATCCCAATCCCATGCAATGG 1050  
 Db 963 CAAGGAGATACCAATGGAACGCAACCGCTAATG-----TCAA 1001  
 Qy 1051 TGATGCTACCACTGTGCTGCTGACATTTTCAACAGCCCGAATGTAGTTAACTTGAGCAGGAC 1110  
 Db 1002 CGGACAGCCTCAGTCGCGGATATCTTCAATAGCGAATGTTGTGAACTTGAGCAGGCA 1061  
 Qy 1111 TGGATTTGAGAAATTTGGCAAAATTTGGAGGACCGGAAAGAAACCATGCTGTGTGCTTTA 1170  
 Db 1062 AGGATTTGAGAAATTTGATGAATTTAGAGAAACCGGAAAGAGCTTGGATCGTTGTCTTTA 1121  
 Qy 1171 TGCACCATGTGCCCCCTACTGCCAGGCTATGAGGAATCTTATGTTCACTTAGCAGACAA 1230  
 Db 1122 TGACCGGTGTGCCCCGTTTGTCAAGCAATGGAAGCTTCAATCGATGAATTTGGCGGATAA 1181  
 Qy 1231 GTTAGCAGGCTCAACAGGATGAAGGTTTGGAAATTTAGAGCAGATGAGCAAGAAAGA 1290  
 Db 1182 GCTGGGAGG---AAGTGTGTGAAGTTGCTAAGTTTAGAGCTGATGTCACCAAGGA 1238  
 Qy 1291 ATTTGCAAGAGTGAATGCAATTTGGGAAGCTTCCCTTACGATATATTTTCCCAAGCA 1350  
 Db 1239 CTTTGCCAAAAGAGTTGCAAGCTTGGAGCTTTCCCAACAACTACTGTGTTCCTCCAAAGAA 1298  
 Qy 1351 TTGCTCTCGCCACATAAAGTATCCTCAGAAAGAGAGATGTTCAATTTCTTGATGGC 1410  
 Db 1299 CTCTTCAAG---ACCAATCAAGTATCCATCAGAGAAGAGAGATGTTGTTCTTTGACATC 1355  
 Qy 1411 ATTTGTAAT 1420  
 Db 1356 TTTCTGAAT 1365  
 RESULT 12  
 AY062665  
 LOCUS Arabidopsis thaliana PRH26 protein (At4g21990; Fln20.90) mRNA, complete cds.  
 DEFINITION  
 ACCESSION AY062665  
 VERSION AY062665.1 GI:17065177  
 KEYWORDS FLI CDNA.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
 REFERENCE 1 (bases 1 to 1654)  
 AUTHORS Southwick,A., Karlin-Neumann,G., Nguyen,M., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Bath,J., Hayashizaki,Y., Ishida,J., Kamiya,A., Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K., Sakano,H., Satou,M., Seki,M., Shinn,P., Yamada,K., Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.  
 Direct Submission  
 Submitted (14-NOV-2001) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA  
 COMMENT e-mail for correspondence: arab@sequence.stanford.edu  
 RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL CDNA : 'RIKEN Arabidopsis Full-length cDNA'); Seki,M., Narusaka,M., Ishida,J.,

Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.  
 The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Nguyen,M., Palm,C.J., Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Bath,J., Chen,H., Cheuk,R., Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P., Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Southwick,A. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W. (SSP/Stanford) contributed equally to this work as PIs.

# FEATURES

## source

1. 1654  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="IV"  
 /clone="RAFLO8-16-K20"  
 /ecotype="Columbia"  
 /notes="This clone is in pBluescript"

## gene

1. 1654  
 /genes="At4g21990; Fln20.90"  
 77. 1453  
 /genes="At4g21990; Fln20.90"  
 /codon\_start=1  
 /product="PRH26 protein"  
 /protein\_id="AALJ2743.1"  
 /db\_xref="GI:17065178"

## CDS

/translation="MALAINVSSSSSAISSSPSSDLKVTIKIGSLRLNRTNVSAA  
 SLSGKRSVSKALNVQSITKESIVASEVTEKLDVVEDEFEELAKLENASPLEIKH  
 KALEKFNDAIAFSGAEDVALIEYALHTGRPVRFSLDTRLNPETRYLFDVTEKH  
 GIRIEYMPDPAVEAOALVRNKLGLFSFVEDGHQSCCRKVRPLRLALKGLRAWITGGR  
 KOQSPCTREIPIVQVDPVEGLDGGVSLVKNPVANVENDWFLRMDVDPNTL  
 HAQYVSGCEPCTRAVLPGHREGRWEDAKCEGLHKGNIKENTNGNATANVN  
 GTASVADIFNSENVNLRSQIENLMKLENKEAWIVLYAPMCPFCQAMBSFDELA  
 DKLGSGGVKAFKADGDDQDFAKELQLGSPFTILVFPKNSRPRIKYPSEKRDVDSL  
 TSPNLNLR"

## ORIGIN

Query Match 40.3%; Score 657.2; DB 15; Length 1654;  
 Best Local Similarity 69.5%; Pred. No. 3,2e-162;  
 Matches 966; Conservative 0; Mismatches 388; Indels 36; Gaps 4;

Qy 31 AATGGCCCTCGCTTTCACCTTCTCAATTTCCGACCACTTCCACCTTCCCATCATCGGA 90  
 Db 88 AATCAACGTTTCTTCATCTTCTTCTTCTGCGATCTCAAGCTCTAGCTTCCCTTCTTCA 147  
 Qy 91 ACCCAAACTTCCGCAAAATTTGGGTCAATTTAGGATTTTCGAGAGGCCCAATTCGAGCGCGT 150  
 Db 148 TCTCAAAAGTAACAAAATCGGATCATTTAGGTTATTGAATCGTACCAATGCTCTGCGGC 207  
 Qy 151 TAATTTCAATTTATCTCAAGAGCGAGCTTGGTAAAGCCCGTTAAACGCCCAACCTCCACG 210  
 Db 208 TTCTCTGAGTTTGTCCGGGAAGAGATCTCCGTTGAAAGCTCTTAATGTGCAATCAATTAC 267  
 Qy 211 CAGGATTCATTTGTTCTCTCCGAGCAACCACTGTTGCTTCTCTCTGAGAGCAA 270  
 Db 268 AAGGAATTCATTTGTTCTTCTGAGGTTACAGAGAAGCTAG-----ATGTGGTGA 318  
 Qy 271 AGAGGAAGATTTTGAACAGATACCGATGATCTCGACAATGCTTCACTCTCTGAAATCAT 330  
 Db 319 AGTTGAAGACTTTGAGGAAGTACGAAGAGATTAGAGATGCTTCTCTCTGAGATCAT 378  
 Qy 331 GGATAGAGCCCTCGACAAATTCGGCAACGACATAGCTATTGCTTCAAGTGGTGTGGAAGA 390  
 Db 379 CGATAAAGCTCTTGAAGATTTGGGAATGACATTGCTATTGCTTATGCTTAGTGAGCTGAAGA 438  
 Qy 391 TGTGCTCTTTCATGATGATGCGAAATTTGACGGTTCGACCTTTAGGGTTTTCAGTTTGA 450  
 Db 439 TGTGCTCTGATGATGATGCTCAATTTAACTGGAAGACCTTATAGGGTATTTCAGTTTGA 498

```
QY 451 CACTGGGAGACTGAACCCAGAACTTATCAACTTTTGTATGCGGTTGAGAAGCAATTATGG 510
D 499 TACAGGAGATTTGAATCCAGAAACATATAGACTCTTCGATACCTGGAGAAACATTACGG 558
QY 511 AATTTCGATTTGAGTACATGTTCCCTGATGCTGTTGAGGTTTCAGGCATTTGGTGAGGAGTAA 570
D 559 TATTTCGATTTGAGTATATGTTCTCTGATGCTGTTGAGGTTTCAGGTTTCAAGCTTTAGTTAGAACAA 618
QY 571 GGGGTTATTCCTTTCTACGAGGATGGGCCAACAGAGTGTTCAGGGTGAGAAAGGTGAG 630
D 619 GGGTTGTTCTCTTCTTATGAAGATGGTTCACAGAGTGTTCGCGGTATTTAGAAAGGTGAG 678
QY 631 GCCTTTAAGGAGGGCCCTTAAGGCTCTCAGAGCATGATCACTGTCAGAGGAAAGACCA 690
D 679 ACCTTTGAGGCGTCCGTTGAAGGGTTTACGTGCTTGGATCACTGTCAGAGGAAAGATCA 738
QY 691 GTCACTCTGTTACTAGGCTCTGAATACCGGTTGTTTTCAGGTTGATCCGGCTTTTGAAGGAAT 750
D 739 ATCACTCCGGACAGATCGGAGATTCCTGTTGTTTTCAGGTTGATCCAGTGTGTTGAAGGAT 798
QY 751 GGATGGTGAATTTGGAAGCTTGGTGAAGTGAACCCCTGTTGCAAAATGTGAAGGCCATGA 810
D 799 AGATGGTGGAGTTGTTGTTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 858
QY 811 CATATGGAACCTCTTAGGACCATGATGTCCTGTGAATTCCTTGCATGCAAAAGGATA 870
D 859 TGTGTTGAATTTCTTGAAGGACGATGGATGTTCTGTTCAACACGCTTACGCTCGGGTTA 918
QY 871 TGTGTTTCAATTCGAGTGTGAGCCCTGACCTAGGAGCTGTTTACCTGGGCAACATGAAGGGA 930
D 919 TGTGTTTCAATCGATGTGAGCCATGACGAGAGCGGTTTTCGGGTCAGCATGAGAGAGA 978
QY 931 AGGAGGTGTTGTTGGGAGGATGCAAGAGCTTAAGGAATGTTGCTTTCACAAAGGAATGT 990
D 979 AGGAGATGTTGTTGGGAGGATGCTAAGGCTTAAGAGATGTTGGCTTTCACAAAGGGAAT 1038
QY 991 AAAGCAGCAGAAAGGAGGATGTTTAAGGAATGGGCTATCCCAATCCCATGCAAAATGG 1050
D 1039 CAAGGAGAAATACCAATGGAACCGCAACCGCTTAATG-----TCAA 1077
QY 1051 TGATGCTACCACTGTGCTGACATTTTCAACAGCCGCAATGTAGTTTAACTTTGAGCAGGAC 1110
D 1078 CGGACAGCCTCAGTCCGGATATCTTCAATAGCGAGATGTTGTGAATTTGAGCAGGCA 1137
QY 1111 TGGAAATGAGAAATTTGGCAAAATTTGAGAGGACCGAAAGCAACATGCTGTTGCTTTA 1170
D 1138 AGGATTTGAGAAATTTGATGAAGTTAGAAACCGGAAAGAGGCTTGGATCGTTGCTTTA 1197
QY 1171 TGCACCATGTTGCTGCTTACTGCGGCTATGAGGAAATCTTATGTTGACTTTAGCAGACAA 1230
D 1198 TGCACCGTGGTCCCGTGTGTCAGGCAATGGAAGCTTCAATCGATGAATTTGGCGGATA 1257
QY 1231 GTTAGCAGGTCACACAGGATGAAGGTTTGGAAATTTAGAGCAGATGAGAGCAACAGAA 1290
D 1258 CTGCGGAGG---AAGTGTGTGAAGTTGCTAAGTTTAGAGCTGATGTCAGCAGAGGA 1314
QY 1291 ATTGCAAGAGTGAATGCAATTTGGGAAGCTTCCCTACGATATATATTTTCCCAAGCA 1350
D 1315 CTTTGCCAAAAGGAGTTGCACTTGGAGCTTTCACCAATACATCTGTTGTTCCCAAGAA 1374
QY 1351 TTGCTCTCGGCCAACAAATTAAGTATCCCTCAGAAAGAGAGATGTTGATTTCTTCATGCG 1410
D 1375 CTCCTTCAAG---ACCAATCAAGTATCCATCAGAAAGAGAGATGTTGATTTCTTGCATC 1431
QY 1411 ATTTTGAAT 1420
D 1432 TTTCTTGAAT 1441
```

```
RESULT 13
AY087591
LOCUS Arabidopsis thaliana clone 1702 bp mRNA linear PLN 14-APR-2003
DEFINITION Arabidopsis thaliana clone 36866 mRNA, complete sequence.
```

```
ACCESSION AY087591.1
VERSION GI:21406329
KEYWORDS FULL CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 1702)
AUTHORS Haas,B.J., Volfovsky,N., Town,C.D., Troupkan,M., Alexandrov,N.,
Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
TITLE Full-length messenger RNA sequences greatly improve genome
annotation
JOURNAL Genome Biol. 3 (6), RESEARCH0029 (2002)
PUBMED 12093376
REFERENCE 2 (bases 1 to 1702)
AUTHORS Brover,V., Troupkan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
TITLE Full-length cDNA from Arabidopsis thaliana
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1702)
AUTHORS Brover,V., Troupkan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
TITLE Direct Submission
JOURNAL Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
Malibu, CA 90265, USA
COMMENT This clone sequence is one of 5,000 Ceres full-length cDNAs made
available to TIGR and Genbank. The following quality assessment of
this set was done by comparison with known proteins: two percent of
the clones are estimated to be 5'-truncated; less than one percent
are 3'-truncated; approximately two percent represent alternative
splice variants, including unspliced introns and spliced exons; one
percent may contain premature stop codons; five percent may have
frame shifts in a coding region. A sequence is considered to be
5'-truncated if it lacks the translation initiation start (ATG). A
C-terminal end of the encoded protein. Please note that these cDNA
sequences are derived from the Ws or Laer ecotypes and therefore
may contain polymorphisms when compared to sequences from Col-0.
Genset carried out the library production and sequencing of the
full-length clones. Ceres, Inc. carried out the clustering of the
5' sequences, selection of clones, and sequence assembly.
```

FEATURES

source

```
1. 1702
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="36866"
78. 1454
/codon_start=1
/product="PRH26 protein"
/protein_id="AAM5133.1"
/db_xref="GI:21593184"
/translation="MALAINVSSSSSSAIISSSPFSDLKVTKIGSLRLNRTNVSA
SLKSGKRSVYKALNVOSITKESIVASEVTEKLJDEVEFEELAKLENASPLIMD
KALKFGNDIAIFSGAEDVALIEVHLTGPRVYFSLDTGRLENPEYRLFDVTEVKH
GIRLEYMPDPAVEQVALVRNKGFSFYEDGHCCEIRKVRPLRALKGLRAWITGQR
KQSPGTRSEIPEVQVDFVFGDGGSLVKWNPVANVEGNDVNFRTWMDVPNTL
HAAGVSTGSEPCRAVLPGQHEREGRWWDEDAKAKCEGLHKGNIKENTNGNATANVN
GTATVADIFNSENVNLSRQGIENLMKLENKEAWIVVLYAPWCPFCQAMEASFDELA
DKLGSGGVKAKFRADGQDQFAKKELQSGFPTILVFPKNSRPPIKYPKSRDVSLS
TSFLNLVR"
```

CDS

ORIGIN

```
Query Match 40.3%; Score 657.2; DB 15; Length 1702;
Best Local Similarity 69.5%; Pred. No. 3.2e-162;
Matches 966; Conservative 0; Mismatches 388; Indels 36; Gaps 4;
QY 31 AATGGCCCTCGCTTTCACCTTCTTCAATTTCCGACCAACTCCACCTCCCATCGCATCGGA 90
D 89 AATCAACGTTTCTTCATCTTCTTCTTCTGCGATCTCAAGCTAGTCTTCTTCTTCTTCA 148
QY 91 ACCCAAACTTCGCAAAATTTGGTCAATTAGGATTTTGGAGAGGCCCATTTGGAGGCGCCGT 150
```

Db 149 TCTCAAGATACAAAAATCGGATCATTTAGGTTATTTGAATCGTACCAATGCTCTCGCGC 208  
 Qy 151 TAAATTTCAATTTATCTCAAGACCGGAGCTTGGTAAAGCCCGTTAAACGCCGCAACCTCCACG 210  
 Db 209 TTCTCTGATTTGTCGGGAAGAGATCCTCGTGAAGCTCTTAATGTGCAATCAATTAC 268  
 Qy 211 CAAGGATTCATTTGCTCTCGCAGCAACCAATCGTTGCTCTGCTCTGAGACGAA 270  
 Db 269 AAAGGAATCCATTTGCTCTGAGGTTACAGAAAGCTAG-----ATGTGTGGA 319  
 Qy 271 AGAGGAGATTTTCAACAGATAGCCAGTGATCTCGAATGCTTCACTCTTGAATCAT 330  
 Db 320 AGTTGAAGACTTTGAGGAACCTAGCAAGAGATTAGAAATGCTTCTCTCTTGAGATCAT 379  
 Qy 331 GGATAGAGCCCTCGACAAATTCGCAACGACATAGCTATTGCTTCACTCTTGAATCAT 390  
 Db 380 GGATAAGACTCTTGAGAAATTTGGGAATGACATGCTATTGCTTTAGTGAGCTGAAGA 439  
 Qy 391 TGTGCTTTGATGATGCGAAATTAAGCGGTCGACCCCTTTAGGCTTTTCAGTTTGA 450  
 Db 440 TGTGCTCTGATGATGCTCATTTAACTGGAAGACCTTATAGGCTATTTCAGTTTGA 499  
 Qy 451 CACTGGGAGCTGAACCCAGAACTTATCACTTTTGTGATGCGGTTGAGAGCAATTATGG 510  
 Db 500 TACAGGAGATTGAATCCAGAAACATATAGACTCTTCGATACCGTGGAGAAACATTACGG 559  
 Qy 511 AATTGCGATTGAGTACATGTTCCCTGATGCTGTTGAGGTTCCAGGCATTTGGTGAGGATA 570  
 Db 560 TATTGCGATTGAGTATGTTCTGATGCTGTTGAGGTTCAAGCTTTAGTAGGAACAA 619  
 Qy 571 GGGGTTATCTCTTTACAGGATGGCCACCAAGAGTTGTGAGGTTGAGAAAGGTGAG 630  
 Db 620 GGGTTGTTCTCTTCTATGAAGATGTCACCAAGAGTTGTCGCTATTAGAAAGGTGAG 679  
 Qy 631 GCCTTTAAGGAGGCCCTTAAGGCTCTCAGAGCATGATTAAGTCTCAGAGGAAGACCA 690  
 Db 680 ACCTTTGAGGCGTGTGAGGAGTTTACGCTGCTGATCACTGCTCAGAGGAAGATCA 739  
 Qy 691 GTCACTGCTACTAGTCTGAAATACCGGTTGTTGAGGTTGATCCGGCTTTTGAGGGAAT 750  
 Db 740 ATCAACCGGCAAGATCGGAGATCTCTGTTGTTGAGGTTGATCCAGTGTGAGAGATT 799  
 Qy 751 GGAATGTTGGAATTTGGAAGCTTGGTGAAGTGAACCCCTGTTGCAATGTGAAGGCCATGA 810  
 Db 800 AGATGGTGGAGTTGGTATGTTGGTGAAGTGAATCCGGTTGCTAATGTTGAAGGGAATGA 859  
 Qy 811 CATATGAACCTCTTAGGACCATGATGTCCTGCTGTAATTCCTTGATGCAAAAGGATA 870  
 Db 860 TGTGGAATTTCTTGAGGACGATGGATGTTCTCTGTCAACACGCTTCACGCTGCGGGTTA 919  
 Qy 871 TGTTCATTTGGGTGTCAGCCCTGCTAGCTAGCCCTGTTTACCTGGGCAACATGAAGGGA 930  
 Db 920 TGTTTCAATCGGATGTCAGCATGTCAGAGCGGTTTTCGCGGTGTCAGATGAGAGA 979  
 Qy 931 AGGAGGTGTTGGTGGAGGATGTCAAAGCTTAAGGAATGGTCTTCAACAAAGGAATGT 990  
 Db 980 AGGAGATGTTGGTGGGAAGATGCTAAGGCTAAAGAGTGGGCTTCAACAAAGGAATAT 1039  
 Qy 991 AAAGCAGCAAGAGAGAGATGTTAATGAATGGGCTATCCCAATCCCATGCAATGG 1050  
 Db 1040 CAAGGAGAAATACCAATGGAACCGAACCCGCTAATG-----TCAA 1078  
 Qy 1051 TGATGCTPACCATGTCCTGCACATTTTCAACAGCCCGAATGTAGTTAACTTGAGCAGGAC 1110  
 Db 1079 CGGACAGCCACAGCTCGGATATCTTCAATAGCGAAGATGTTGTGAACCTTGAGCAGGCA 1138  
 Qy 1111 TGGAAATTGAAATTTGCAAAATTTGAGGAGCCGAAAGAACCAATGGCTGTTGCTGCTTTA 1170  
 Db 1139 AGGATTGAGAAATTTGATGAATTAGAGAACCGGAAGAGGCTTGGATCGTTGCTGCTTTA 1198  
 Qy 1171 TGACCATGTTGCCCTTACTGCGAGGCTATGAGGAATCTTATGTTGAGTTAGCAGACAA 1230

Db 1199 TGCACCGTGTGCGCCGTTTTTGTCTAGGCAATGGAAGCTTCATTCGATGAATGGCGGATAA 1258  
 Qy 1231 GTTAGCAGGCTCAACAGGATGAAGTTTGGAAAATTTAGACAGATGAGAACAGAAAGA 1290  
 Db 1259 GCTGGGAGG--AAGTGTGTGAAGTTGCTAAGTTTAGAGCTGATGGTGACCAAGAGA 1315  
 Qy 1291 ATTTGCAAGAGTGAACTGCAATTTGGGAAGCTTCCCTACGATATATATTTTCCCAAGCA 1350  
 Db 1316 CTTTGCCAAAAGAGGTTGCAGCTTGGAAGCTTTCCAAATACTGTGTGTTCCCAAGAA 1375  
 Qy 1351 TTGCTCTCGCCCAACATAAAGTATCCCTCAGAAAAGAGAGATTTGATTCCTTGATGGC 1410  
 Db 1376 CTCTTCAAG--ACCAATCAAGTATCCATCAGAAAGAGAGATGTTGATTTTGACATC 1432  
 Qy 1411 ATTTGTAAT 1420  
 Db 1433 TTCTTGAAT 1442

RESULT 14  
AF428445

LOCUS 1709 bp mRNA linear PLN 18-OCT-2001  
 DEFINITION Arabidopsis thaliana AT4g21990/Fin20\_90 mRNA, complete cds.  
 ACCESSION AF428445  
 VERSION AF428445.1 GI:16226613  
 KEYWORDS FLI CDNA.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana

REFERENCE  
AUTHORS

1 (bases 1 to 1709)  
 Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,  
 Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,  
 Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,  
 Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,  
 Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,  
 Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,  
 Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,  
 Davis,R.W., Theologis,A. and Ecker,J.R.  
 Arabidopsis cDNA clones  
 Unpublished

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

2 (bases 1 to 1709)  
 Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,  
 Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,  
 Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,  
 Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,  
 Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,  
 Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,  
 Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,  
 Davis,R.W., Theologis,A. and Ecker,J.R.  
 Direct Submission

TITLE  
JOURNAL

Submitted (05-OCT-2001) Salk Institute Genomic Analysis Laboratory  
 (Signal), Plant Biology Laboratory, The Salk Institute for  
 Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,  
 USA

## COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the  
 collection and clustering of AFLP cDNAs (RAFL cDNA : 'RIKEN  
 Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J.,  
 Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,  
 Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGECC (SSP) Consortium members carried out the  
 sequencing and annotation of the RAFL cDNAs: Cheuk,R., Chen,H.,  
 Kim,C.J., Koesema,E., Meyers,M.C., Shim,P., Banh,J., Bowser,L.,  
 Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G.,  
 Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M.,  
 Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A.,  
 Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S.,  
 Davis,R.W., Theologis,A., and Ecker,J.R.

Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to  
 this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)







```

Db      1147 AGGATTGAGAAATTGATGAAGTTAGAGAACCGGAAGAGGCTTGGATCGTTGTGCTTTA 1206
Qy      1171 TGCACCATGGTGGCCCTTACTGCCAGGCTATGGAGGAATCTTATGTTGACTTAGCAGACAA 1230
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1207 TGCACCGTGGTGGCCGTTTGTGAGGCAATGGAAGCTTCATTCGATGAAATTGGCGGATAA 1266
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      1231 GTTAGCAGGTCACACAGGATGAAGTTGGAATAATTTAGAGCAGATCGAGAACAGAAAGA 1290
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1267 GCTGGAGG---AAGTGGTGTGAAGTTGCTAAGTTTAGAGCTGATGGTGACCCAGAAAGGA 1323
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      1291 ATTTGCAAGAGTGAACCTGCAATTGGGAAGCTTCCCTACGATATTATTTTCCCAAAGCA 1350
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1324 CTTTGCCTAAAGAGTTGAGCTTGAAGCTTCCAACTACTTGTGTTCCCAAGAA 1383
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      1351 TTGCTCTCGGCCAACAAATAAGTATCCCTCAGAAAGAGAGATGTTGATTCCTTTGATGGC 1410
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1384 CTCCTCAAG---ACCAATCAAGTATCCATCAGAGAAGAGAGATGTTGATTCTTTGACATC 1440
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      1411 ATTTGTAAAT 1420
        |||||
Db      1441 TTTCTTGAAT 1450
        |||||

```

Search completed: February 20, 2006, 23:16:02  
 Job time : 8433 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 20:40:11 ; Search time 941 Seconds  
(without alignments)

11537.497 Million cell updates/sec

Title: US-10-731-525-7

Perfect score: 1629

Sequence: 1 gcacgaggagagaccacata.....taaaaaaaaaaaaaaaaaaa 1629

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_21.\*

1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002as.\*  
7: Geneseqn2002bs.\*  
8: Geneseqn2003as.\*  
9: Geneseqn2003bs.\*  
10: Geneseqn2003cs.\*  
11: Geneseqn2003ds.\*  
12: Geneseqn2004as.\*  
13: Geneseqn2004bs.\*  
14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1629	100.0	1629	3	Az50066 Soybean a
2	916.8	56.3	1730	13	Adx13202 Plant ful
3	904	55.5	1795	3	Az50065 Soybean a
4	657.2	40.3	1377	3	Aac43415 Arabidops
5	657.2	40.3	1702	3	Aac38867 Arabidops
6	656.6	40.3	1712	3	Aac47286 Arabidops
7	655	40.2	1714	3	Aac37691 Arabidops
8	651.6	40.0	1398	6	Abz12693 Arabidops
9	651.6	40.0	1398	6	Adg88184 A. thalia
10	651.6	40.0	1398	8	Ada68112 Arabidops
11	651.6	40.0	1398	12	Adn74398 thale cre
12	588.2	36.1	1687	3	Aac32729 Arabidops
13	578.4	35.5	1685	3	Aac48356 Arabidops
14	577.4	35.4	1365	6	Abz14500 Arabidops
15	577.4	35.4	1365	6	Adg88214 A. thalia
16	577.4	35.4	1365	6	Adg88215 A. thalia
17	577.4	35.4	1365	12	Adn73974 Thale cre
18	568	34.9	110000	3	Continuation (4 of
19	538.2	33.0	1827	3	Aaz50067 Wheat ade

20	512.6	31.5	1840	13	ADT17458	Adt17458 Plant cdn
21	467.4	28.7	1380	4	AAH42329	Aah42329 Nucleotid
22	435.2	26.7	1210	3	Aaz50064	Aaz50064 Impatiens
23	364.4	22.4	1156	13	Adx29365	Adx29365 Plant ful
24	364.2	22.4	1215	3	Aaz50063	Aaz50063 Corn aden
25	356.6	21.9	1183	13	Adx32480	Adx32480 Plant ful
c 26	353.8	21.7	568	9	ACL21376	ACL21376 DNA clone
c 27	284	17.4	422	9	ACL21373	ACL21373 DNA clone
c 28	277.4	17.0	623	13	ACN56923	ACN56923 Cotton gy
29	271.4	16.7	1031	3	AAC52097	AAC52097 Arabidops
c 30	270	16.6	554	9	ACL21377	ACL21377 DNA clone
31	266.6	16.4	471	12	ADF93046	Adp93046 Cotton ex
32	256.4	15.7	585	13	ACN49050	ACN49050 Cotton pr
c 33	247.4	15.2	670	9	ACL21371	ACL21371 DNA clone
c 34	223.8	13.7	588	13	ACN49122	ACN49122 Cotton pr
c 35	221.6	13.6	475	9	ACL21375	ACL21375 DNA clone
c 36	211.6	13.0	552	13	ACN56966	ACN56966 Cotton gy
c 37	199.6	12.3	783	9	ADA29776	Ada29776 DNA encod
c 38	194.8	12.0	441	9	ACL21374	ACL21374 DNA clone
c 39	159	9.8	705	13	ADS61808	Ads61808 Bacterial
c 40	152.6	9.4	251	13	ADT06475	Adt06475 Nicotiana
c 41	149.4	9.2	810	11	ABD15340	Abd15340 Pseudomon
c 42	148.8	9.1	659	9	ACL21370	ACL21370 DNA clone
c 43	141.4	8.7	514	10	AD881371	Ad881371 Arabidops
c 44	126	7.7	614	9	ACL21378	ACL21378 DNA clone
c 45	126	7.7	624	9	ACL21385	ACL21385 DNA clone

#### ALIGNMENTS

RESULT 1  
AAZ50066  
ID AAZ50066 standard; cdna; 1629 BP.  
XX AAZ50066;  
AC AAZ50066;  
XX  
DT 04-MAY-2000 (first entry)  
XX  
DE Soybean adenosine-5'-phosphosulfate reductase cdna clone s12.pk0064.g4.  
XX  
KW Adenosine-5'-phosphosulfate reductase; APS reductase; soybean;  
XX sulphate assimilation protein; SAP; clone s12.pk0064.g4; ss.  
XX  
OS Glycine max.  
XX  
XX  
FH Key Location/Qualifiers  
FT CDS 32..1432  
FT /\*tag= a  
FT /product= "Soybean APS reductase"  
XX  
XX WO200004161-A1.  
XX  
XX 27-JAN-2000.  
XX  
XX 13-JUL-1999; 99WO-US015808.  
XX  
XX 14-JUL-1998; 98US-0092833P.  
XX  
XX (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX  
XX Falco SC, Allen SM;  
XX WPI; 2000-182427/16.  
XX P-PSDB; AAZ50066.  
XX  
XX Nucleic acid fragments encoding sulfate assimilation proteins in plants  
XX and seeds useful as probes for isolating cDNAs and genes encoding  
XX homologous proteins, and in producing transgenic plants.  
XX  
XX Claim 3; Page 34-35; 47pp; English.  
XX  
XX The present cdna sequence encodes soybean adenosine-5'-phosphosulfate

CC (APS) reductase, a sulphate assimilation protein (SAP). The cDNA clone is  
CC designated s12.pk0064.94, isolated from soybean two week old developing  
CC seedlings treated with 2.5ppm chlorimuron. APS reductase polypeptides are  
CC useful for producing antibodies for screening cDNA expression libraries  
CC to isolate full length APS reductase cDNA clones. Chimeric genes  
CC comprising the APS reductase gene when transformed in a host cell can  
CC alter the level of SAP expression. This can be used to increase sulphur  
CC containing compounds in the cell, including the nutritionally critical  
CC amino acids cysteine and methionine  
XX  
SQ Sequence 1629 BP; 461 A; 309 C; 420 G; 439 T; 0 U; 0 Other;

Query Match 100.0%; Score 1629; DB 3; Length 1629;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1629; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCACGAGAGAGAACCCATACAGCTAGTTAATGGCCCTCGCTTTCACCTTCTTCAATTC 60  
Db 1 GCACGAGAGAGAACCCATACAGCTAGTTAATGGCCCTCGCTTTCACCTTCTTCAATTC 60

Qy 61 CGCACCAACTTCCACCTTCCCATCATCGGAACCCAAACTTCCGCAAAATGGGTCAAATAG 120  
Db 61 CGCACCAACTTCCACCTTCCCATCATCGGAACCCAAACTTCCGCAAAATGGGTCAAATAG 120

Qy 121 GATTCGGAGAGGCCCATTCGAGGCGCGTTAATTTCAATTTATCTCAAGACGAGCTT 180  
Db 121 GATTCGGAGAGGCCCATTCGAGGCGCGTTAATTTCAATTTATCTCAAGACGAGCTT 180

Qy 181 GGTAAAGCCGTTAAACGCCGAACCTCCACGCAAGGATTCATTTCTCTCGCAGCAAC 240  
Db 181 GGTAAAGCCGTTAAACGCCGAACCTCCACGCAAGGATTCATTTCTCTCGCAGCAAC 240

Qy 241 AACCATCGTCTTCTGCTTCTGAGACGAAAGAGAGATTTGAACAGATAGCAGTGA 300  
Db 241 AACCATCGTCTTCTGCTTCTGAGACGAAAGAGAGATTTGAACAGATAGCAGTGA 300

Qy 301 TCTCGACATGCTTCACTTCTGAAATCATGGATAGAGCCCTCGACAAATTCGCGACGA 360  
Db 301 TCTCGACATGCTTCACTTCTGAAATCATGGATAGAGCCCTCGACAAATTCGCGACGA 360

Qy 361 CATAGCTATTCGCTTCACTTCTGAAATCATGGATAGAGCCCTCGACAAATTCGCGACGA 420  
Db 361 CATAGCTATTCGCTTCACTTCTGAAATCATGGATAGAGCCCTCGACAAATTCGCGACGA 420

Qy 421 GGGTCGACCCCTTTAGGGTTTTCAGTTTGGACACTGGGAGACTGAACCCAGAAACTTATCA 480  
Db 421 GGGTCGACCCCTTTAGGGTTTTCAGTTTGGACACTGGGAGACTGAACCCAGAAACTTATCA 480

Qy 481 ACTTTTTCATCGGTTGAGAGCATATGGAATTCGCATTCGATTCGATGTTCCCTGATGC 540  
Db 481 ACTTTTTCATCGGTTGAGAGCATATGGAATTCGCATTCGATTCGATGTTCCCTGATGC 540

Qy 541 TGTTCAGGTTTCAGGCTTTCAGGAGTAAGGGGTTATTTCTTCTACGAGGATGGGCA 600  
Db 541 TGTTCAGGTTTCAGGCTTTCAGGAGTAAGGGGTTATTTCTTCTACGAGGATGGGCA 600

Qy 601 CCAAGAGTTTTCAGGGTTGAGAAAGGTGAGGCTTTTAAAGGAGGGCCCTTAAAGGGTCTCAG 660  
Db 601 CCAAGAGTTTTCAGGGTTGAGAAAGGTGAGGCTTTTAAAGGAGGGCCCTTAAAGGGTCTCAG 660

Qy 661 AGCATGGATTAATGCTCAGAGAAAGACCACTGCTGCTAGTCTGGAATACCGGT 720  
Db 661 AGCATGGATTAATGCTCAGAGAAAGACCACTGCTGCTAGTCTGGAATACCGGT 720

Qy 721 TGTTCAGGTTTCAGGCTTTCAGGAGTAAGGGGTTATTTCTTCTACGAGGATGGGCA 780  
Db 721 TGTTCAGGTTTCAGGCTTTCAGGAGTAAGGGGTTATTTCTTCTACGAGGATGGGCA 780

Qy 781 GAACCCCTGTTGCAATGTGAAGGGCCCATGACATATGGAACCTTCTTGAACCATCAATGT 840  
Db 781 GAACCCCTGTTGCAATGTGAAGGGCCCATGACATATGGAACCTTCTTGAACCATCAATGT 840

Qy 841 GCCTGTGAATTCCTTGCATGCAAAAGGATATGTTTCCATTGGGTGTGAGCCCTGCACTAG 900

Db 841 GCCTGTGAATTCCTTGCATGCAAAAGGATATGTTTCCATTGGGTGTGAGCCCTGCACTAG 900

Qy 901 GCCTGTTCCTTACCTGGGCAACATGAAGGGAAGGAGGTGTGTGGAGAGATGCCAAGC 960  
Db 901 GCCTGTTCCTTACCTGGGCAACATGAAGGGAAGGAGGTGTGTGGAGAGATGCCAAGC 960

Qy 961 TAAGGAATGTGGTCTTTCACAAAGGAATGTAAAGCAGCAAGAGAGAGGATGTTAATGG 1020  
Db 961 TAAGGAATGTGGTCTTTCACAAAGGAATGTAAAGCAGCAAGAGAGAGGATGTTAATGG 1020

Qy 1021 AAATGGGCTATCCCAATCCCATGGTGTGATGTACCATCTGTGCTGCACTTTTCAA 1080  
Db 1021 AAATGGGCTATCCCAATCCCATGGTGTGATGTACCATCTGTGCTGCACTTTTCAA 1080

Qy 1081 CAGCCCCGAATGTAGTTAACTTTAGCAGGACTGGAATTTGAGAAATTTGGCAAAATTTGGAGGA 1140  
Db 1081 CAGCCCCGAATGTAGTTAACTTTAGCAGGACTGGAATTTGAGAAATTTGGCAAAATTTGGAGGA 1140

Qy 1141 CCGAAGGAACCATGGCTTGTGTGCTTATGCAACCATGGTGCCCTTACTGCCAGGCTAT 1200  
Db 1141 CCGAAGGAACCATGGCTTGTGTGCTTATGCAACCATGGTGCCCTTACTGCCAGGCTAT 1200

Qy 1201 GGAGGAATCTTATGTTGACTTTAGCAGACAAGTTAGCAGGGTCAACAGGGATGAAGTTGG 1260  
Db 1201 GGAGGAATCTTATGTTGACTTTAGCAGACAAGTTAGCAGGGTCAACAGGGATGAAGTTGG 1260

Qy 1261 AAAATTTAGAGCAGATGGAGAACAGAAAGATTTGCAAGAGTGAACTGCAATTTGGGNAAG 1320  
Db 1261 AAAATTTAGAGCAGATGGAGAACAGAAAGATTTGCAAGAGTGAACTGCAATTTGGGNAAG 1320

Qy 1321 CTTCCCTACGATATTTATTTTCCCAAGCATTCGTCGCGCAACAATAAAGTATCCCTC 1380  
Db 1321 CTTCCCTACGATATTTATTTTCCCAAGCATTCGTCGCGCAACAATAAAGTATCCCTC 1380

Qy 1381 AGAAAGAGAGATGTTGATTCCTTTGATGGCAATTTGTAATGCCCTTAAGATGAGGATATCA 1440  
Db 1381 AGAAAGAGAGATGTTGATTCCTTTGATGGCAATTTGTAATGCCCTTAAGATGAGGATATCA 1440

Qy 1441 GGAATTTTCTTCTTTTGGGTTGCAATTCACCTTTGACTATACGTACAGCGGGTTCT 1500  
Db 1441 GGAATTTTCTTCTTCTTTTGGGTTGCAATTCACCTTTGACTATACGTACAGCGGGTTCT 1500

Qy 1501 TCTTTATGCTTATACGTGTATATACCATTCGTTTACAGATTCCTTCTGTAACCTGTTGGA 1560  
Db 1501 TCTTTATGCTTATACGTGTATATACCATTCGTTTACAGATTCCTTCTGTAACCTGTTGGA 1560

Qy 1561 AGTGGGAATGAGGTTTATACAAATAAGATACCTCAGTTTTGAATGGTTTTAAAAA 1620  
Db 1561 AGTGGGAATGAGGTTTATACAAATAAGATACCTCAGTTTTGAATGGTTTTAAAAA 1620

Qy 1621 AAAAAAAAA 1629  
Db 1621 AAAAAAAAA 1629

RESULT 2  
ADX13202  
ID ADX13202 standard; cDNA; 1730 BP.  
XX  
AC ADX13202;  
XX  
DT 21-APR-2005 (first entry)  
XX  
Plant full length insert polynucleotide seqid 7777.  
XX  
DE  
KW plant protectant; plant growth regulant; gene therapy; plant;  
KW recombinant DNA construct; physical array; plant breeding marker;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;  
KW extreme osmotic condition; pathogen tolerance; pest tolerance;  
KW growth rate; cell cycle pathway; disease resistance;  
KW galactomannan production; lignin production; plant growth regulator;  
KW yield; plant growth; plant development; seed oil; protein yield;







Db	1320	AGCAGAAGGAATATGCAAGAGTGAACTGCAGTTGGGAAGCTTCCCCACAATACTTCTCT	1379
Qy	1341	TCCCAAGAGCATTCCTCTCGGCCCAACAATAAGTATCCCTCAGAAAAGAGAGATGTTGATT	1400
Db	1380	TCCCCAAGCACTCTTCT--CAACCAATTAAGTACCTTCAGAAAAGAGAGATGTTGATT	1436
Qy	1401	CCTTGATGGCATTGTAAATGCCCTTAAGATGAGGAT	1436
Db	1437	CATTGACGGCATTGCGAATGCCCTTACGGTGATGGT	1472
RESULT 4			
AAC43415			
ID	AAC43415 standard; DNA; 1377 BP.		
XX	AC		
AC	AAC43415;		
XX	XX		
DT	17-OCT-2000 (first entry)		
XX	XX		
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 39173.		
XX	XX		
KW	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway; metabolic pathway;		
KW	promoter; termination sequence; ss.		
XX	XX		
OS	Arabidopsis thaliana.		
XX	XX		
PN	EP1033405-A2.		
XX	XX		
PD	06-SEP-2000.		
XX	XX		
PF	25-FEB-2000; 2000EP-00301439.		
XX	XX		
PR	25-FEB-1999;	99US-01218225P.	99US-0138540B.
PR	05-MAR-1999;	99US-0123180P.	99US-0138847E.
PR	09-MAR-1999;	99US-0123548P.	99US-0139119P.
PR	23-MAR-1999;	99US-0125788P.	99US-0139452P.
PR	25-MAR-1999;	99US-0126264P.	99US-0139492P.
PR	28-MAR-1999;	99US-0126785P.	99US-0139456P.
PR	01-APR-1999;	99US-0127462P.	99US-0139457P.
PR	06-APR-1999;	99US-0128234P.	99US-0139458P.
PR	08-APR-1999;	99US-0128714P.	99US-0139459P.
PR	18-APR-1999;	99US-0129845P.	99US-0139460P.
PR	19-APR-1999;	99US-0130077P.	99US-0139461P.
PR	21-APR-1999;	99US-0130449P.	99US-0139459P.
PR	23-APR-1999;	99US-0130510P.	99US-0139457P.
PR	28-APR-1999;	99US-0130891P.	99US-0139454P.
PR	30-APR-1999;	99US-0132048P.	99US-0139455P.
PR	30-APR-1999;	99US-0132407P.	99US-0139456P.
PR	04-MAY-1999;	99US-0132484P.	99US-0139457P.
PR	06-MAY-1999;	99US-0132485P.	99US-0139458P.
PR	06-MAY-1999;	99US-0132486P.	99US-0139459P.
PR	07-MAY-1999;	99US-0132487P.	99US-0139456P.
PR	11-MAY-1999;	99US-0132863P.	99US-0139457P.
PR	14-MAY-1999;	99US-0134218P.	99US-0139458P.
PR	14-MAY-1999;	99US-0134219P.	99US-0139459P.
PR	14-MAY-1999;	99US-0134422P.	99US-0139460P.
PR	18-MAY-1999;	99US-0134370P.	99US-0139459P.
PR	19-MAY-1999;	99US-0134768P.	99US-0139456P.
PR	20-MAY-1999;	99US-0134941P.	99US-0139457P.
PR	21-MAY-1999;	99US-0135124P.	99US-0139458P.
PR	24-MAY-1999;	99US-0135353P.	99US-0139459P.
PR	25-MAY-1999;	99US-0135829P.	99US-0139460P.
PR	27-MAY-1999;	99US-0136021P.	99US-0139459P.
PR	28-MAY-1999;	99US-0136392P.	99US-0139456P.
PR	01-JUN-1999;	99US-0136782P.	99US-0139457P.
PR	03-JUN-1999;	99US-0137222P.	99US-0139458P.
PR	04-JUN-1999;	99US-0137528P.	99US-0139459P.
PR	07-JUN-1999;	99US-0137502P.	99US-0139456P.
PR	08-JUN-1999;	99US-0137724P.	99US-0139457P.
PR	08-JUN-1999;	99US-0138094P.	99US-0139458P.
PR	10-JUN-1999;	99US-0140055P.	99US-0139459P.
PR	10-JUN-1999;	99US-0140085P.	99US-0139460P.
PR	10-JUN-1999;	99US-0140325P.	99US-0139459P.
PR	10-JUN-1999;	99US-0140331P.	99US-0139457P.
PR	10-JUN-1999;	99US-0140332P.	99US-0139458P.
PR	10-JUN-1999;	99US-0140333P.	99US-0139459P.
PR	10-JUN-1999;	99US-0140334P.	



```
Db 1062 AGGATTGAGAAATTGATGAAGTTAGAGAACCGGAAAGAGCGTTGGATCGTTGTGCTTTA 1121
Qy 1171 TGCAACCATGGTGCCTTACTCCAGGCTATGAGAGAAATCTTATGTTCGACTTAGCAGACAA 1230
Db 1122 TGCACCGTGTGTCGGTGTTCCTGTCAGGCAATGGAAGCTTCATTCGATGAATGGCGGATAA 1181
Qy 1231 GTTAGCAGGTCAACAGGATGAAGTTGGAATAATTTAGCAGATGAGAACAGAAAGA 1290
Db 1182 GCTGGGAGG---AAGTGGTGTGAAGGTTGCTTAAGTTTAGAGCTGATGGTGACCAAGAAGGA 1238
Qy 1291 ATTTGCAAGAGTGAACATGCAATGGGAAGCTTCCTACGATATATTTTTTCCCAAGCA 1350
Db 1239 CTTTGCCAAAAGAGTTGAGCTTGGAGCTTTCCAAATATCTTGTTTCCCAAGAA 1298
Qy 1351 TTGCTCTCGGCCAACAAATAAGTATCCCTCAGAAAAGAGATGTTGATTCCTTTGATGGC 1410
Db 1299 CTCCTTCAAG---ACCAATCAAGTATCCATCAGAGAGAGAGATGTTGATTCCTTTGACATC 1355
Qy 1411 ATTTGTAAT 1420
Db 1356 TTTCTTGAAT 1365

RESULT 5
AAC38867
ID AAC38867 standard; DNA; 1702 BP.
XX
AC AAC38867;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 22518.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 14-JUN-1999; 99US-0138847P.
PR 16-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140895P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143824P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
```

```
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147039P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158023P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159323P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.

PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161921P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 40.3%; Score 657.2; DB 3; Length 1702;
Best Local Similarity 69.5%; Pred. No. 1.6e-178;
Matches 966; Conservative 0; Mismatches 398; Indels 36; Gaps 4;

Qy 31 AATGGCCCTCGCTTCACTTCTCAATTCGCGACCAACTTCCACCTTCCCATCATCGGA 90
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 91 ACCCAAACTTCGCGAAATTTGGTCAATTAGATTTCCGAGAGGCCCATTCGAGGGCGGT 150
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 149 TCTCAAAAGTAACAAAAATCGGATCATTTGAGGTTATTGAATCGTACCAATGCTCTCGGCG 208
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 151 TAATTTCAATTTATCTCAAAGACGGAGCTTGGTAAAGCCGTTAAACGCCGAACCTCCACG 210
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 209 TTCTCTGAGTTTGTCCGGAGAGATCTCTCGTGAAGAGCTCTTAATGTGCAATCAATTAC 268
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 211 CAAGGATTTCCATTTGTTCTCTCGCAGCAACCACTCGTTGCTTCTGCTCTGAGACGAA 270
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 269 AAAGGAATCCATTGTTGCTTCTGAGGTTACAGAGAAGCTAG-----ATGTGGTGA 319
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 271 AGAGGAAGATTTTGAACAGATAGCCAGTATCTCGCAACTGCTTCACTCTTGAATCAT 330
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 320 AGTTGAAGACTTTGAGGAACTTAGCAAGAGATTAGAGAAATGCTTCTCTCTTGAGATCAT 379
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 331 GGATAGAGCCCTCGACAAATTTCCGCAACGACATAGCTATTGCCCTTCAGTGGTGTGAAGA 390
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 380 GGATAAAGCTCTTGAGAAAGTTTGGGAATGACATTTGCTATTGCTTTAGTGAGCTGAAGA 439
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 391 TGTGCTTTGATTGAGTATGCGAAATTTGACGGGTCGACCCCTTTAGGGTTTTCAGTTTGA 450
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 440 TGTGCTCTGATTGAGTATGCTCATTTAACTGGAAGACCTTATAGGGTATTTCAGTTTGA 499
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 451 CACTGGGAGACTGAACCCAGAAAATTATCAACTTTTGTATCGGTTGAGAGCATTATGG 510
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 500 TACAGGGAGATTGAATCCAGAAAACATATAGACTCTTCGATACCCGTGGAGAAACATTACGG 559
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 511 AATTCCGATTCAGTACATGTTCCCTGATGCTGTTGAGGTTCCAGGCATTGGTGAGAGTAA 570
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 560 TATTCGGATTGAGTATATGTTTCTGATGCTGTTGAGGTTCAAGCTTTAGTTAGGAACAA 619
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 571 GGGGTTATTTCTTTTCTACGAGGATGGGCACCAAGAGTGTTCGAGGGTGAGAAAGGTGAG 630
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 620 GGGTTTGTCTCTTTCTATGAAGATGGTCAACAGAGTGTTCGCCGTATTAGAAAGGTGAG 679
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 631 GCCTTTAAGGAGGGCCCTTAAGGGTCTCAGAGCATGGAATACTGTCAGAGGAAGAACCA 690
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 680 ACCTTTGAGCGGTGCGTTGAAAGGGTTTACGTGCTTGGATCACTGGTCAGAGAAAGATCA 739
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 691 GTCACCTGCTACTAGTCTGAAATACCGGTTGTTTCAGGTTCGATCCGGCTTTTTCAGGGGAT 750
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 740 ATCACCCGGGCAAGATCGGAGATTCCTGTTGTTTCAGGTTGATCCAGTGTTCGAGGATT 799
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 751 GGATGGTGGAAATTTGGAAGCTTTGGTGAAGTGGAAACCTGTTGCAAAATGGAAGGGCATGA 810
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 800 AGATGGTCGAGTTGGTAGTTTGGTGAAGTGGAAATCCCGGTTCGTAATGTTGAAGGGAATCA 859
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 811 CATATGGAACCTTCCTTAGGACCAATGATGCTGCTGTAATTCCTTTCATGCAAAAGGATA 870
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 860 TGTTTGGAAATTTCTTGAGGACGATGGATGTTTCTCTCAACACGCTTCAGCTGCGGGTTA 919
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 871 TGTTCATTTGGGTGTGAGCCCTGACCTAGGCTGCTGTTTACCTGGGCAACATGAAGGGA 930
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 920 TGTTCATTCGATGTGAGCCCATGCGAGAGCGGTTTTCGCGGGTCAGCATGAGAGAGA 979
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 931 AGGGAGGTGGTGGTGGGAGGATGCCAAAGCTAAGGAAATGTGTCTTTCACAAAGGAATGT 990
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 980 AGGAGAGTGGTGGGGAAGATGCTAAGGCTAAAGAGTGTGGCCCTTCACAAAGGGAATAT 1039  
Qy 991 AAACAGCAGAAAAGAGGAGGATGTTAATGGAAATGGGCTATCCCAATCCCATGGCAAAATGG 1050  
Db 1040 CAAGGAGATACCAATGGAAACGCAACCGCTAATG-----TCAA 1078  
Qy 1051 TGATGCTACCACTGTGCTGACATTTTCAACAGCCCGAATGTAGTTAACTTTGACGAGGAC 1110  
Db 1079 CGGACAGCCACAGTCGGCGGATATCTTCAATAGCGAAGATGTTGTGAATCTTTGACGAGCA 1138  
Qy 1111 TGGAAATTGAGAAATTGGCAAAATGGAGGACCGAAAGCAATGCGCTTGTGTGCTTTA 1170  
Db 1139 AGGAAATTGAGAAATTGAGAAATGAGAAATGAGAAACCGAAAGAGCGCTTGGATCGTTGTGCTTTA 1198  
Qy 1171 TGCAACCATGTGTCCTCTACTGCCAGGCTATGGAGGAATCTTATGTTGACTTTAGCAGCAA 1230  
Db 1199 TGCAACCGTGTGTCCTGTTTGTGAGGCAATGGAAGCTTTCATCGAATGGAATGGCGGATAA 1258  
Qy 1231 GTTAGCAGGTCACAGGAGTGAAGGTTGGAAATTTAGACGAGATGGAGAACAGAAAGA 1290  
Db 1259 GCTGGAGG---AAGTGTGTGAAGTTGCTAAGTTTAGAGCTGATGTCAGCAGAAAGGA 1315  
Qy 1291 ATTTGCAAGAGTGAATGCAATGGCAATGGGAAGCTTCCCTACGATATATTTTCCCAAGCA 1350  
Db 1316 CTTTGCCTGAGGAGTTGCAAGCTTGGAAAGCTTTCCCAACATACTGTGTGTTCCCAAGAA 1375  
Qy 1351 TTCTGTCTCGGCCAACAAATGAAGTATCCCTCAGAAAAGAGAGATGTCGATTCCTTGTATGGC 1410  
Db 1376 CTCTTCAAG---ACCAATCAAGTATCCATCAGAGAGAGAGAGTGTGATCTTTTGACATC 1432  
Qy 1411 ATTTGTAAT 1420  
Db 1433 TTTCTTGAAT 1442

RESULT 6  
AAC47286  
ID AAC47286 standard; DNA; 1712 BP.  
XX  
AC AAC47286;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 53259.  
XX  
XX Hybridisation assay; generic mapping; gene expression control;  
XX protein identification; signal transduction pathway; metabolic pathway;  
XX promoter; termination sequence; ss.  
XX  
XX Arabidopsis thaliana.  
XX  
XX EPI033405-A2.  
XX  
XX  
PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-00301439.  
XX  
XX 25-FEB-1999; 99US-012825P.  
XX 05-MAR-1999; 99US-0123180P.  
XX 09-MAR-1999; 99US-0123548P.  
XX 23-MAR-1999; 99US-0125788P.  
XX 25-MAR-1999; 99US-0126264P.  
XX 29-MAR-1999; 99US-0126785P.  
XX 01-APR-1999; 99US-0127462P.  
XX 06-APR-1999; 99US-0128234P.  
XX 08-APR-1999; 99US-0128714P.  
XX 16-APR-1999; 99US-0129845P.  
XX 19-APR-1999; 99US-0130077P.  
XX 21-APR-1999; 99US-0130449P.  
XX 23-APR-1999; 99US-0130510P.  
XX 23-APR-1999; 99US-0130891P.  
XX 30-APR-1999; 99US-0131449P.  
XX 99US-0132048P.

PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134376P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 14-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139452P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140895P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 02-JUL-1999; 99US-0142154P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.

PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	23-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153707P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159333P.
PR	14-OCT-1999;	99US-0159637P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	22-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 40.3%; Score 656.6; DB 3; Length 1712;		
Best Local Similarity 70.8%; Pred. No. 2.4e-178;		
Matches 974; Conservative 0; Mismatches 359; Indels 42; Gaps 6;		
Qy	83	TCATCGGAACCCAAACTTCCGCAAAATTGGGTCAATTAGGATTTCCGAGAGCCCATGGA 142
Db	133	TCATTGGAGCCAAAGTTTCGCAAAATTGGTTCGTTGAGGTTATTGGATCGTGTTCATGTT 192
Qy	143	GGCGCGGTAAATTCAATTATCTCAAGACGG---AGCTTGGTAAAGCCCGTTAACGCC 199
Db	193	GCTCCTGTCTCTGAAATCTATCTGGGAAGCATCATCATCTGTAAACCTTTAAACGCT 252
Qy	200	GAACCTCCAGCAAGGATTCATTTCTCTCCAGCAACCAACCATCGTTGCTTCTGCT 259
Db	253	GAACCAAGACAAAGGATTCATGATTCCTCTCGGCAACAACTGTAGAGAAATGCA 312
Qy	260	TCTGAGACGAAGAG-----GAAGATTTTGAACAGATAGCAGTGATCTCGACAAT 310
Db	313	GAGGAAGTTGAAGTGTGAGATTTGAGAGCTTGTGAAGCTTAGAGAAT 372
Qy	311	GCTTCACCTCTTGAATCATGAGTAGAGCCCTGCACAAATTCGGCAACGACATAGCTATT 370
Db	373	GCTTCACCTCTTGAGATTTATGCAAAAGCTCTTGAGAAATACGGGAACGATATGCCATT 432
Qy	371	GCCTTCAGTGCTGCTGAAGATGTTGCTTGAATGAGTATGCAAAATTCACGGGTGACCC 430
Db	433	GCAITTAGTGTGCAGAAAGATGTTCTCTTAITTAGTAGCTCATTTGACTGGGAGGCCA 492
Qy	431	TTTAGGGTTTTAGTTTTGGACACTGGGAGACTGAACCCAGAAACTTATCAACTTTTGTAT 490
Db	493	TTTAGAGTATTTAGTTTGGATACAGGGAGGTTGAATCCTGAGACGTATCGGTTTTTTCGAT 552
Qy	491	CGGGTTGAGAAGCATATGGAATTCGCANTTGAGTACATGTTCCCTGATGCTGTTCAGGTT 550
Db	553	CGGGTGGAGAAGCACATATGGGATTAGGATTTAGGTATGATATATGTTCTCTGATTCGTGAGGTT 612
Qy	551	CAGGCATTTGTGAGGAGTAAGGGTTATCTCTTTCTACGAGGATGGGCCAACAGATGT 610
Db	613	CAAGTTTTGGTTAGGACGAGGATGTTCTCTTTTATGAGGATGGTCAACAGAGTGT 672
Qy	611	TGCAGGGTGAAGAGGTGAGGCCCTTTAAGGAGGGCCCTTAAAGGGTCTCAGAGCATGGATA 670
Db	673	TGCCGTGTTCCGAAGGTGAGACCTTTGAGGGGTGCTCTCAAGGGTTTAAAGGCTTGGATT 732
Qy	671	ACTGTTGAGAGAAAGACAGTCACTCGTACTAGGTCTGAAATACCGGTTGTTTCAGGTT 730
Db	733	ACTGTTGAGAGAAAGATCAATCTCCGGGCAAGGCTGAGATTCGGTTCGTTTCAGGTT 792
Qy	731	GATCCGGCTTTTGGGGAAATGATTCGGAATTTGGAAGCTTGGTGAAGTGAACCCCTGTT 790
Db	793	GATCCGGTGTTTGAAGGTTTGGATGGAGATTGGTAGTTTGGTGAAGTGGAAATCCGGTT 852
Qy	791	GCAATGTGAAGGGCCATGACATATGGAACCTTCCTTAGGACCATGAAATGTGCTGTGAAT 850



```
Db      853  GCGAATGTTGAAGGGAATGATGTTTGGAACTTCTTGAGGACTATGGATGTTCCGGTTAAC 912
Qy      851  TCCTTGCATGCAAAAGGATATGTTTCCATTGGGTGTGAGCCCTGCACCTAGGCCCTGTTTTA 910
Db      913  ACATTGCATCGGCAGGGTATATATCGATTGGATGTGAGCCTTGCAAGAAAGCGGTTTTA 972
Qy      911  CCTGGGCAACATGAAGGAAGGAGGTGGTGGTGGAGGATGCAAAAGCTTAAGGAATGT 970
Db      973  CCGGCTCAGCAGCAGAGAGAGGAGATGGTGGTGGGAAGATGCTAAAGCCCAAGGAATGT 1032
Qy      971  GGTCTTCACAAGCAATGTTAAAGCAGCAGCAAGAGGAGGTGTTAATGGAAATGGGCTA 1030
Db      1033  GGACTTCACAAGGGAATGTCA----AAGAAACTCCGATGATGCTAAAGTGAACGGGAA 1089
Qy      1031  TCCCAATCCCATGCAAAATGGTGATGCTACCACTGTGCTGACATTTTCAACAGGCCCAAT 1090
Db      1090  TCGAAT-----CCGCTGTTGCAGATATCTTTAAGATGAGAT 1128
Qy      1091  GTAGTTAACTTGACAGACTGGAATTGAGAAATTTGCCAAATTTGGAGGACCCGAAAGGAA 1150
Db      1129  CTTGTGACTTTGACGCGCAGGGGATTTGAGAAATTTGATGAATTTGGAGAACCGTAAAGAG 1188
Qy      1151  CCATGGCTTGTGCTTTATGCAACCATGTTGCCCTTACTCCAGGCTATGGAGGAATCT 1210
Db      1189  CTTGGATCGTCGTCTTTATGCTCCGTGGTGGCCCTTTTGTCAAGCCATGGAAGCATCG 1248
Qy      1211  TATGTTGACTTTAGCAGCAACAGTTAGCAGGGTCAACAGCGATGAAGTTTGGAAATTTAGA 1270
Db      1249  TATGATGACTGGGGGATNAATTTGGCTGG---AGTGGGATTAAGTTGCCAATTCAGA 1305
Qy      1271  GCAGATGAGAACAGAAAGAAATTTGCAAGAGTGAATGCAATTTGGGAAGCTTCCCTACG 1330
Db      1306  GCAGATGGTGACCAAGAGGAGTTTGCTTAAGCAGGAATTCAGCTCGGTAGCTTCCCTACC 1365
Qy      1331  ATATTATTTTCCCAAGCATTCGTCTCGGCCCAACATTAAGTATCCCTCAGAAAGAGA 1390
Db      1366  ATCTGTTTTTCCCTAAGAACTCATC---GAGACCGATCAAGTATCCGTCTGGAAGAGA 1422
Qy      1391  GATGTTGATTCCTTGATGGCAATTTGTAATGCCCTTAAGATGAGGATATCAGGAAA 1445
Db      1423  GATGTTGAGTCTTTGACTTCGTTCTGAACTTGTCCGTAAGTAAGTAACCAACAAA 1477

RESULT 7
AAC37691
ID AAC37691 standard; DNA; 1714 BP.
XX
AC AAC37691;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 18317.
XX
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
```

PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144332P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	21-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0147260P.
PR	02-AUG-1999;	99US-0147303P.
PR	02-AUG-1999;	99US-0146388P.
PR	03-AUG-1999;	99US-0146389P.
PR	04-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148117P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151068P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155488P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
<hr/>		
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
<hr/>		
Query Match 40.2%; Score 655; DB 3; Length 1714;		
Best Local Similarity 70.8%; Pred. No. 7e-178;		
Matches 973; Conservative 0; Mismatches 360; Indels 42; Gaps 6;		
Qy	83	TCATCGGNAACCAACATTCGCAAAATGGGTCAATTAGGATTCGGAGAGCCCAATTGGA 142
Db	133	TCATTGGAGCCAAAGTTTCGCAAAATGGTTCGTTGAGGTTATTGGATCGTGTTCATGTT 192
Qy	143	GGCGCCGTTAATTTCAATTTATCTCAAGACGG---AGCTTGGTAAAGCCCGTTAAACGCC 199
Db	193	GCTCCTGTGCTCTGAAATCTATCTGGGAAGCATCATCTGTTAAACCTTTAAACGCT 252
Qy	200	GAACCTCCAGCAAGGATTCATTTGTTCTCTCGAGGCAACCAACCATCGTTGCTTCGCT 259
Db	253	GAACCAAGACAAACGATTTCAATGATTCCTCTCGGCAACAATGGTAGCAAGAAATGCA 312
Qy	260	TCTGAGACGAAGAG-----GAAGATTTGAAACAGATAGCCAGTGATCTCGACAAT 310
Db	313	GAGGAAGTTGAAGTGGTTGAGATTTGAGGATTTTGAAGAGCTTCTAAGAAAGTTAGAGAT 372
Qy	311	GCTTCACCTCTTGAAATCATGGATAGAGCCCTCGACAAATTCGGCAACGACATAGCTATT 370
Db	373	GCTTCACCTCTTGAGATTTATGACAAAGCTCTTGAGAAATACGGGAACGATATCGCCATT 432
Qy	371	GCCTTCAGTGCTGGAAGATGTTGCTTTGATTGATGATGCGAAATTCACCGGTCGACCC 430
Db	433	GCAATTTAGTGTGAGAGAGATGTTGCTCTTATTGAGTACGCTCATTTGACTGGGAGGCCA 492
Qy	431	TTTAGGGTTTTCAGTTTCGACACTGGGAGACTGAACCCAGAAACTTATCAACTTTTGTAT 490
Db	493	TTTAGAGTATTAGTTTGGATACAGGGAGGTTGAATCTCTGAGACGTATCGGTTTTTCGAT 552
Qy	491	GCGGTTGAGAAAGCATTTAGGAATTCGCAATTCAGTACATGTTCCCTGATGCTGTTGAGGTT 550
Db	553	GCGGTGGAGAGCACTATGGGATTAGGATTAGTATATGTTTCTCTGATTTCTGTTGAGGTT 612
Qy	551	CAGGATTTGGTGAGGAGTAAGGGTTATTCTCTTCTACGAGGATGGGCAACCAAGATGT 610
Db	613	CAAGGTTTGGTTAGGAGCAAGGGAATGTTCTCTTTTATAGAGGATGGTCATCAGAGATGT 672
Qy	611	TGCAGGTTGAGAAAGGTGAGGCTTTAAGGAGGGCCCTTAAGGGTCTCAGAGCATGGATA 670
Db	673	TGCCGTGTTTCGAAAGGTGAGACCTTTGAGGCGTGTCTCTCAAGGGTTTAAAGGCTTGGATT 732
Qy	671	ACTGTCAGAGGAAGACCAAGTCACCTCGTACTAGGTCTGAATAACCGGTTGTTTCAGGTT 730

Db 733 ACTGTCAGAGGAAGATCAATCTCCGGGCAAGGCTCGAGATCCGGTTGTCAGGTT 792  
Qy 731 GATCCGGCTTTGAGGGAATGGATGGTGAATTCGAAGCTTGGTGAAGTGAACCCCTGTT 790  
Db 793 GATCCGGTGTGTTGAAGGTTGAGATGGTGGAGTGGTAGTTGGTGAAGTGAATCCGGTT 852  
Qy 791 GCAATGTGAAGGCCCATGACATATGGNACTTCCTTAGGACCATGAATGTGCGCTGTGAAT 850  
Db 853 GCGAATGTGAAGGGAATGATGTTTGGNACTTCTTAGGACTATGATGTTCCGGTTAAC 912  
Qy 851 TCCTTGCATGCAAAAGGATATGTTTCCATTGGTGTGAGCCCTGACCTAGGCTGTTTGA 910  
Db 913 ACATTGATCGCGCAGGATATATATCGATTGATGATGAGCCCTGCACAAAGCGGTTTA 972  
Qy 911 CCTGGCAACATGAAAGGAGGAGGTGGTGGGAGGATGCCAAAGCTTAAGGAATGT 970  
Db 973 CCGGCTCAGCAGCAGAGAGGAGGAGATGGTGGTGGGAAGATGCTAAAGCCCAAGGAATGT 1032  
Qy 971 GGTTCTTCAAAAGGAATGTAAAGCAGCAGAGAGAGGATGTATATGGAATGGGCTA 1030  
Db 1033 GGACTTCAAAAGGGAATGTCA---AAGAAACTCCGATGATGCTAAAGTGAACGGGAA 1089  
Qy 1031 TCCCAATCCCATGCAATGGTGTGCTACCACTGCTGCTGACATTTTCAACAGCCCGAAT 1090  
Db 1090 TCGAAT-----CGCTGTGCGATATCTTTAAGAGTGAGAA 1128  
Qy 1091 GTAGTTAACTTGACAGCACTGGAATTCGAAATTTGCAAAATTTGGAGGACCGAAAGAA 1150  
Db 1129 CTTGTGACTTTGACAGCAGGAGGATGAGNAATTTGATGAATTTGGAGAACCGTAAAGAG 1188  
Qy 1151 CCATGGCTGTGTTTATGCAACATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1210  
Db 1189 CTTGGAATCGTGTGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1248  
Qy 1211 TATGTTCACTTAGCAGCACTGAGGCTGACAGGATGAGGTTGGAATTTAGA 1270  
Db 1249 TATGATGAATGCGGGAATAATGGCTGG---AAGTGGGATTAAGGTTGCCNAATTCAGA 1305  
Qy 1271 GCAGATGGAGAAAGAAAGAAATTTGCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1330  
Db 1306 GCAGATGGTGCACAGAGGATTTGCTAAGCAGGAATTTGAGCTGCTGCTGCTGCTGCTGCT 1365  
Qy 1331 ATATATTTTCCAAAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1390  
Db 1366 ATTCTGTTTTTCCCTAAGAACTCATC---GAGACCGATCAAGTATCGCTCTGAGAGAGA 1422  
Qy 1391 GATGTTGATTTCTTGTGCAATTTGTAATGCTTAAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1445  
Db 1423 GATGTTGATGCTTTGACTTTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1477

RESULT 8

ABZ12693  
ID ABZ12693 standard; DNA; 1398 BP.  
XX AC ABZ12693;  
XX DT 21-JAN-2003 (first entry)  
XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 498.  
XX KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
XX OS Arabidopsis thaliana.  
XX PN W0200216655-A2.  
XX PD 28-FEB-2002.  
XX PF 24-AUG-2001; 2001WO-US026685.  
XX PR 24-AUG-2000; 2000US-0227866P.  
PR 26-JAN-2001; 2001US-0264647P.

PR 22-JUN-2001; 2001US-0300111P.

XX (SCRI ) SCRIPPS RES INST.  
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX Harper JP, Kreps J, Wang X, Zhu T;  
XX WPI; 2002-304127/34.

PT Identifying a stress condition to which a plant cell has been exposed and  
PT producing plants with increased tolerance to these abiotic stresses.

PS Claim 144; SEQ ID NO 498; 577pp + Sequence Listing; English.

XX The invention relates to identifying a stress condition to which a plant  
CC cell has been exposed, comprising: (a) contacting nucleic acid  
CC representative of expressed polynucleotides in the plant cell with an  
CC array or probes representative of the plant cell genome; and (b)  
CC detecting a profile of expressed polynucleotides in the plant cell  
CC characteristic of a stress response. The method is useful in the  
CC production of transgenic plants, cells and seeds and in producing plants  
CC with increased tolerance to abiotic stress. The present sequence is that  
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
CC in methods of the invention. Note: The sequence data for this patent is  
CC not represented in the printed specification but is based on sequence  
CC information supplied to Derwent by the European Patent Office

XX Sequence 1398 BP; 357 A; 219 C; 407 G; 415 T; 0 U; 0 Other;

Query Match 40.0%; Score 651.6; DB 6; Length 1398;  
Best Local Similarity 70.9%; Pred. No. 6.1e-177;  
Matches 966; Conservative 0; Mismatches 354; Indels 42; Gaps 6;

Qy 83 TCATCGAACCACCAATTCGCAAAATGGGTCAATTAGGATTTGGGAGAGCCCATGGA 142  
Db 67 TCATGGAGCCAAAGTTTCGCAAAATGGGTTCGTGAGGTTATGGATCGGTTCATGTT 126  
Qy 143 GGCGCCCTTAATTTCAATTTATCTCAAAAGCGG---AGCTTGGTAAAGCCCGTTAAAGCC 199  
Db 127 GCTCTGTGTCTCGAATCTATCTGGGAAGGATCATCACTGTAAACCTTTAAACGCT 186  
Qy 200 GAACCTCCAGCAAGGATTCATTTGCTCTCGCAGCAACAACCATCGTTGCTTGTCT 259  
Db 187 GAACCAAAAGCAAAAGGATTCATGATTTCTCTTGGCGCAACAATGTTAGCAAAATGCA 246  
Qy 260 TCTGAGAGCAAGAG-----GAGATTTTGAACAGATAGCCAGTCTCGACAAT 310  
Db 247 GAGGAAGTTGAAGTGGTTGAGATTGAGATTGAGAGCTTTGAGAGCTTCTAAGAGTTAGAAAT 306  
Qy 311 GCTTCACCTCTTGAAATCATGGATAGAGCCCTCGACAAATTCGCAACACGACATAGCTATT 370  
Db 307 GCTTCACCTCTTGAGATTATGGACAAAGCTTTGAGAAATACGGGAAGATATCGCCATT 366  
Qy 371 GCCTTCAGTGGTCTGGAAGATGTTGCTTGAATGAGTATCGGAAATTTGACGGTTCGACC 430  
Db 367 GCATTTAGTGGTCAGAAAGATGTTGCTCTTATTTAGTACGCTCATTTGACTGGGAGGCCA 426  
Qy 431 TTTAGGGTTTTCAGTTTGGACACTGGGAGACTGAACCCAGAACTTATCAACTTTTGTAT 490  
Db 427 TTTAGAGTATTTAGTTTGGATA CAGGAGGTTGAATCTCTGAGACGATATCGGTTTTCGAT 486  
Qy 491 GCGGTTGAGAGCAATTTAGAAATTCGCAATTTGAGTACATGTTTCCCTGATGCTGTTGAGGTT 550  
Db 487 GCGGTTGAGAGCACTATGGGATTTAGGATTTAGTATATGTTTCTGATTTCTGTTGAGGTT 546  
Qy 551 CAGGCAATTTGAGGAGTAAAGGGTTATTTCTCTTTTACAGGATGGGCAACCAAGAGTGT 610  
Db 547 CAAGGTTTGGTTAGGAGCAAGGATTTGTTCTCTTTTTATGAGGATGTCATCAGGAGTGT 606  
Qy 611 TGCAGGTTGAGAAAGGTGAGGCCCTTTAAGGAGGCGCTTAAGGCTCTCAGAGCATGATA 670  
Db 607 TGCCGTTGTTCCAAAGGTGAGACCTTTGAGGCGGTCTCTCAAGGGGTTTAAAGGGCTTGGATT 666





```
QY 371 GCCTTCAGTGGTCTGAAGATGTTGCTTTGATTGAGTATGCGAAATTCAGCGGTCGACCC 430
Db 367 GCATTTAGTGGTCAGAAAGATGTTGCTCTTATTAGTACGCTCATTTGACTGGAGGCCA 426
QY 431 TTTAGGGTTTTAGTTGGACACTGGGAGACTGAACCCAGAAACTTATCAACTTTTTTGAT 490
Db 427 TTTAGAGTATTAGTTGGATACAGGGAGGTTGAATCTTGAGAGTATCGGTTTTTCGAT 486
QY 491 GCGGTTGAGAGCAATATGGAATTCGCAATTGAGTACATGTTCCCTGATGCTGTTGAGGTT 550
Db 487 GCGGTGAGAGCACTATGCGATTAGGATTAGGATATGATTTCTCTGATTTCTGTTGAGGTT 546
QY 551 CAGGCATTTGTCAGAGTAAAGGTTATTCTCTTTCTACGAGGATGGCCACCAAGAGTGT 610
Db 547 CAAGGTTTGGTTAGGACCAAGGATTTGTTCTCTTTTATGAGGATGGTCTACGAGAGTGT 606
QY 611 TGCAGGGTGCAAAAGGTTGAGGCCCTTTAAGGAGGCGCTTTAAGGGTCTCTCAGAGCATGGATA 670
Db 607 TGCCTGTTTGAAGGTTGAGACCTTTGAGGCGTGTCTCAAGGGTTTAAAGGCTTGGATT 666
QY 671 ACTGGTCAGAGGAAGACCACTGACCTGGTACTAGTCTGAAATACCGGTTGTTTCAGGTT 730
Db 667 ACTGGTCAGAGGAAGATCAATCTCCGGGCAAGGCTGAGATTCGGGTTGTTTCAGGTT 726
QY 731 GATCCGGCTTTGAGGGAATGATGGTGGATTTGGAAGCTTGTGGAAGTGAACCCGTTT 790
Db 727 GATCCGGGTTTGAAGGTTTGGATGGTGGAGTTGGTAGTTTGGTGAAGTGAATCCCGGTT 786
QY 791 GCAATGTGAAGGGCCATGACATATGGAATCTCTTAGGACCATGAATGTCCTGTGTAAT 850
Db 787 GCGAATGTTGAAGGGAATGATGTTTGGAACTTCTTGAGGACTATGGATGTTCCGGTTAAC 846
QY 851 TCCTTGCATCAAAAGGATATGTTTCCATTCGGGTGTAGCCCTGCATAGCCCTGTTTGA 910
Db 847 ACATTTGCATCGGCAGGGTATATATCGATTGGATGTGAGCCTTGCCACGAAGCGTTTTA 906
QY 911 CCTGGGCAACATGAAAGGAAGGAGTGGTGGTGGAGGATGCGCAAGCTTAAGGAATGT 970
Db 907 CCGGTCAGCAGCAGAGAGAGGAGATGGTGGTGGGAAGATGCTTAAAGCCCAAGGAATGT 966
QY 971 GGTCTTCAAAAGGAATTAAGCAGCAGAGAGAGGAGGATGTTTAATGGAATTCGGGCTA 1030
Db 967 GGACTTTCAAAAGGGAATGTCA---AAGAAACTCCGATGATGCTTAAAGTGAACGGGAA 1023
QY 1031 TCCCAATCCCATGCAAAATGGTGATGCTACACTGTGCTGCAATTTTCAACAGCCCGAAT 1090
Db 1024 TCGAAT-----CCGCTGTTGCAGATATCTTTAAGAGTGAAGAT 1062
QY 1091 GTAGTTAACTTGACGAGCACTGGAATTCGAAATTTGGCAAAATTCGAGGACCGAAGAA 1150
Db 1063 CTTGTGACTTTGACGAGCGCAGGGGATGAGAAATTTGATGAAGTTGGGAACCGTAAAGAG 1122
QY 1151 CCATGGCTTGTGCTTTATGCAACATGGTGCCTTACTGCGAGGCTATGGAGGAATCT 1210
Db 1123 CTTUGATGCTGCTGCTTTATGCTCCGTGGTGGCCCTTTTGTCAAGCCATGGAAGCATCG 1182
QY 1211 TATGTTGACTTAGCAGACAAGTTAGCAGGTCACACAGGATGAAGGTTGGAATAATTAGA 1270
Db 1183 TATGATGAACTCGCGGATAAATTTGCTGG---AAGTGGGATTAAGGTTGCCAAATTCAGA 1239
QY 1271 GCAGATGGAAGAACAGAAAGAAATTTGCAAGAGTGAATTCGAATTTGGGAGCTTCCCTACG 1330
Db 1240 GCAGATGGTGACCAAGAGAGTGTGCTAAGCAGGAATTTGCAGCTCGGTAGCTTCCCTACC 1299
QY 1331 ATATTATTATTTCCCAAGACATTCGTCTCGGCCAACAAATAAGTATCCCTCAGAAAAAGAGA 1390
Db 1300 ATCTGGTTTTCCCTAAGAACTCATC---GAGACCGATCAAGTATCCGCTCAGAGAGAGA 1356
QY 1391 GATGTTGATTCCTTGATGSCATTTGTAAATGCTTAAAGATGA 1432
Db 1357 GATGTTGAGTCTTTGACTTCGTTCTTGTGAATCTTGTGCCGATAA 1398
```

## RESULT 11

ADN74398

ID ADN74398 standard; cDNA; 1398 BP.

XX

AC ADN74398;

DT 15-JUL-2004 (first entry)

XX Thale cress cDNA repressed in E2Fa/Dpa expressing plants SeqID 2293.

XX gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;

XX growth regulator; animal feed product; thale cress;

XX cell wall biosynthesis; nitrogen metabolism; carbon metabolism.

XX Arabidopsis thaliana.

OS WO2004035798-A2.

FN 29-APR-2004.

XX 20-OCT-2003; 2003WO-EP011658.

PF 18-OCT-2002; 2002EP-00079408.

XX (CROP-) CROPDESIGN NV.

PA Inze D, De Veylder L, Vlieghe K;

XX WPI; 2004-348466/32.

DR P-PSDB; ADN74399.

XX Altering plant characteristics, useful for producing plants for enzyme or pharmaceutical production comprises modifying in a plant, expression of one or more nucleic acids and/or modifying level or activity of one or more proteins.

XX Claim 1; SEQ ID NO 2293; 134pp; English.

XX This invention relates to a novel method for altering one or more plant characteristics. Specifically, it refers to identifying genes that are up - or down-regulated in transgenic plants overexpressing the heterodimeric E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to alter plant characteristics accordingly. The present invention describes generating transgenic plants for the production of growth regulators, enzymes, therapeutics, pharmaceuticals and animal feed products, where the altered plant characteristics are selected from increased yield or biomass, enhanced survival capacity, stress tolerance, plant architecture or physiology, altered endoreplication and/or altered photosynthesis, transduction, storage lipid mobilisation and/or altered photosynthesis, each relative to the corresponding wild type plants. Accordingly, these sequences can also be useful as positive or negative selectable markers during transformation of cells or tissues. The identified genes play a role in a variety of biological processes such as DNA replication, cell wall biosynthesis, nitrogen and/ or carbon metabolism or they function as transcription factors. This polynucleotide sequence is thale cress cDNA repressed 1.3 fold or more in plants overexpressing the E2Fa/Dpa transcription factor, given in an exemplification of the invention.

XX SQ Sequence 1398 BP; 357 A; 219 C; 407 G; 415 T; 0 U; 0 Other;

Query Match 40.0%; Score 651.6; DB 12; Length 1398;

Best Local Similarity 70.9%; Pred. No. 6.1e-177;

Matches 966; Conservative 0; Mismatches 354; Indels 42; Gaps 6;

QY 83 TCATCGGAACCCAAACTTCGCAAAATTCGGTCAATTAGGATTCGGAGAGGCCCATGGA 142

Db 67 TCATTGGAGCCAAAAGTTTCGCAAAATTCGTTGAGGTTATTCGATCGTTCATGTT 126

QY 143 GCGCGCGTTAATTTCAATTTATCTCAAGACGG---AGCTTGGTAAAGCCGTTAACGCC 199

Db 127 GCTCCTGTCTCTGAATCTATCTGGGAAGCGGATCATCATCTGTGTAAACCTTTAAACGT 186

QY 200 GAACCTCCACGCAAGGATTCATTTCTCTCTCGCAGCAACCAACCATCGTTGTTCTGCT 259



Db	187	GAACCAAGACA	AAAGGATTCAATGATTTCCCTCTGGCGCAACAATGGTAGAGAAATTGCA	246
Qy	260	TCTGAGACGAAAGAG	-----GAAGATTTTGAAACAGATAGCCAGTGATCTCGACAAT	310
Db	247	GAGGAAGTTG	AGTGGTTGAGATTGAGGATTTTGNAGAGCTTGCTAAGAAGTTAGAGAT	306
Qy	311	GCTTCACTCTTGAAATCA	TGGATAGAGCCCTCGACAAATTCGGCAACGACATAGCTATT	370
Db	307	GCTTCACTCTTGAGATTAT	TGGACAAAGCTCTTGAGAAATACGGGAACGATATCGCCATT	366
Qy	371	GCCTTCAGTGGTCTGAAGATGTTGCTTTGATTTGAGTATCGGAAATGACGGGTGACCC	430	
Db	367	GCATTTAGTGGTGAGAAGATGTTGCTCTATTATGAGTACGCTCATTTTGA	CTGGGAGGCCA	426
Qy	431	TTTAGGGTTTTACGTTTTGGGACAC	TGGGAGACTGAACCCAGAGAACTTATCAACTTTTTTGAT	490
Db	427	TTTAGAGTATTTAGTTTGGATACAGGAGGTTGNATCCTGAGAGTATCGGTTTTTCGAT	486	
Qy	491	CGGTTGAGAAAGCAATATGGAATTCGCAATGAGTACATGTTCCCTGATGCTGTGAGGTT	550	
Db	487	CGGTTGAGAAAGCACTATGGCATTAGGATTCAGTATATGTTTCTGATTTCTGTGAGGTT	546	
Qy	551	CAGGCATTTGGTGAGGATTAAGGGTTATTCTCTTTTACGAGGATGGGACCAAGAGTGT	610	
Db	547	CAAGGTTTGGTTAGGACCAAGGATTCGTTCTCTTTTATGAGGATGGTCA	TCAGGAGTGT	606
Qy	611	TGCAGGTTGAAAGGTTGAGGCCCTTTAAGGAGGGCCCTTAAGGCTTCAGAGCATGGATA	670	
Db	607	TGCCGTGTTGCAAAAGTTGAGACCTTTGAGGGTGCTCTCAAGGGTTTAAAGGCTTGAT	666	
Qy	671	ACTGGTCAGAGGAAAGACCACTGCTGTAAGTCTGAAATPACCGGTTGTCAGGTT	730	
Db	667	ACTGGTCAGAGGAAAGATCAATCTCCGGGACAAAGGTCGAGATTCGGTTGTCAGGTT	726	
Qy	731	GATCCGGCTTTGAGGGAATGGATGGTGGAAATTCGGAAGCTTTGGTGAAGTGGAAACCTGTT	790	
Db	727	GATCCGGTGTTTGAAGGTTTTGGATGGTGGAGTTGGTATGTTGGTGAAGTGGAAATCCGGTT	786	
Qy	791	GCAAACTGTAAGGGCCATGACATATGCACTTCCCTTAGGACCATGAATGTCCTGTGAAT	850	
Db	787	CGGAATGTTGAAGGAATGATGTTTGGACCTTTTGAGGACTATGGATGTTCCGGTTAAC	846	
Qy	851	TCCTTGATGCAAAAGGATATGTTTCCATTTGGGTGTGAGCCCTGCACTAGGCCCTGTTTTA	910	
Db	847	ACATTCGATCGGCAGGATATATATCGATTGGATGTCAGCTTCGACGAAGCGGTTTTA	906	
Qy	911	CTCGGCAACATGAAAGGGAAGGAGTGTTGGTGGGAGGATGCAAAAGCTAAGGAATGT	970	
Db	907	CCGGTCAGCACGAGAGAAGGAGATGGTGGTGGGAAGATGCTTAAGCCAAAGGAATGT	966	
Qy	971	GGTCTTCACAAAGGAATGTAAAGCAGCAGAAAGAGGAGGATGTTAATGAAATGGGCTA	1030	
Db	967	GGACTTTCACAAAGGAATGTCA---AAGAAACTCCGATGATGCTAAGTGAACCGGGAA	1023	
Qy	1031	TCCCAATCCCATGCAAAATGGTGATGTACCACTGTGCTGACATTTTTCAACAGCCCGAAT	1090	
Db	1024	TCGAAAT-----CCGCTGTTGCAGATATCTTTAAGATGAGAT	1062	
Qy	1091	GTAGTTAACTTGACACGACTGGAATTTGAGAAATTTGGCAAAATTTGGAGGACCGAAAGGAA	1150	
Db	1063	CTTGTGACTTTGAGCAGCGAGGGGATTTGAGAAATTTGATGAAGTTGGAGAACCGTAAGAG	1122	
Qy	1151	CCATGGCTTGTGTGCTTTATGACCACTGTCGCCCTACTGCCAGGCTATGAGGATCT	1210	
Db	1123	CTTTGGATCTGTGCTCTTATGCTCCCGTGGGCCCTTTTGTCAAGCCATGGGAAGCATCG	1182	
Qy	1211	TATGTTGACTTAGCAGACAAGTTAGCAGGGTCAACAGGGATGAAGTTGAAATTTAGA	1270	
Db	1183	TATGATGAACTGGCGGATTAATTGCTGG---AAGTGGGATTAAGGTTGCCAATTCAGA	1239	
Qy	1271	GCAGATGGAGAACAAAGAAATTTGCAAAAGATGAACTGCAATTTGGGAAGCTTCCCTPACG	1330	

PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139753P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140931P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142330P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 19-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.

PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 23-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 36.1%; Score 588.2; DB 3; Length 1687;  
Best Local Similarity 69.4%; Pred. No. 1.3e-158;  
Matches 895; Conservative 0; Mismatches 343; Indels 51; Gaps 5;  
QY 157 CAATTTATCTCAAGACGGAGCTTGGTAAAGCCCGTTAAAGCCGA-----ACCTCCACG 210

Db 217 CCATTTGCTCAGAGAGGTTACTCTATGAAACCCCTTAAACGCTGAGTCATCACATACAAG 276  
Qy 211 CAAGGATTTCCATTGTTCTCTCGCAGCAACAACCATCGTTGCTTCTCTCTGAGA--- 266  
Db 277 AAGCGATCTGGGTTCTTCTGCTTCACTCTATTTGCTCTCTGAGTTGAAGAAAGA 336  
Qy 267 --CGAAAGAGAGATTTTGAACAGATAGCAGATGATCTCGACAATGCTTCACTCTTGA 324  
Db 337 AGGAGAAAGTTCAAGACTTTGAGCAACTTGTCTAAAGAGCTTGAAGATGCTTCTCCACTGA 396  
Qy 325 AATCATCGATAGACCCCTCGACAAATTCGGCAAGCATAGCTATTGCTTCACTGCTGTC 384  
Db 397 AATCATGGATAAAGCTCTTGAGAAATTTGAGAGCAAAATCGCAATTCCTTTTATGAGGAGC 456  
Qy 385 TGAAGATGTTGCTTTGATGATGCGAAATTTGACCGGTCGACCCCTTAGGGGTTTTTCAG 444  
Db 457 TGAAGATGTTGCAATGATGATGCTGCTTTAACTGGAAGCCATTTAGGGTTTTTAG 516  
Qy 445 TTTGGACACTGGGAGACTGAACCCAGAAATTTATCAACTTTTGTGATCGGTTGAGAAGCA 504  
Db 517 TTTAGATACAGGGAGATTAACCCGGAAGCTACAGGCTCTTTGACGAGTCGAGAGCA 576  
Qy 505 TTATGGAATTCGATTCAGTACATGTTCCCTGATGCTGTTGAGGTTGAGGATGTTGAG 564  
Db 577 GTACGGGATTCGAATTCGATGATGTTTCTGATGCAAGTTGAGGTTCAAGCTTTAGTGAG 636  
Qy 565 GAGTAAGGGGTTATTCCTTTCTACGAGGATGGCCACCAAGAGTGTTCAGGGTGAAGAA 624  
Db 637 GAACAGGGGTTGTTCTATCTATGAAGATGTCATCAAGAGTGTTCGCGTGTGAGGAA 696  
Qy 625 GGTGAGCCCTTTAAGGAGGCCCCTTAAGGGTCTCAGAGCATGGAATACTGGTCAGAGGAA 684  
Db 697 AGTTAGACCTTTGCGTCTGCTCTTAAGGGTCTTAAAGCTTGGATTAACAGACAGAGGAA 756  
Qy 685 AGACCATGACCTGTTACTAGGTTGAAATACCGGTTGTTGAGGTTGATCCGCTTTTGA 744  
Db 757 AGACCAATCTCCGGTACGAGATCTGAGATCTGAGATCTCCCTATTGTTTCAAGTTGATGTTGA 816  
Qy 745 GCGAATGATGTTGGAATTCGAAGCTTGGTGAAGTGAACCCCTGTTGCAAAATGTGAAGGG 804  
Db 817 AGGGTTAGATGGCGGTTGGAAGTCTTGTGAAGTGAATCCCTTTGGCTAATGTTGAAGG 876  
Qy 805 CCATGACATATGGAATCTCTTAGGACCATGAATGTCCTGTGAATTCCTTGTGATGCATA 864  
Db 877 AGCTGATGTGTGAACCTTTCTGAGAACTATGGAATGTTCCGGTGAATGCAATTCATGCACA 936  
Qy 865 AGGATATGTTTCAATGGGTTGAGCCCTGCACTAGGCTGTTTACCTGGGCAACATGA 924  
Db 937 AGGATATGTGTCAATCGGTTGTGAGCCGTGTACTAGGCTGTGCTTCAGGTCAGCATGA 996  
Qy 925 AAGGGAAGGAGGTGTTGGTGGAGGATGCCAAGCTTAAGGAATGTGCTTTCACAAAGG 984  
Db 997 AAGAGAAGGAAGGTGTTGGTGGGAAGATGCTAAGCTTAAAGAAATGTGCTTACACAAAGG 1056  
Qy 985 AAATGTAAAGCAGCAGAAAGAGGAGGATGTTAATGGAAATGGGCTATCCCAATCCCATGC 1044  
Db 1057 GAACATCAAGGAGGAAGTGTG-----CTGC 1083  
Qy 1045 AAATGGTATGCTACCACTGCTGCTGATCATTTTCAACAGCCGAAATGTTAGTTAACTTTGAG 1104  
Db 1084 AGACTCAAGCCCTGCTGCTGCAAGAGATATTTGAAGCAACAATGTGTTGCTTCATTTGAG 1143  
Qy 1105 CAGGACTGGAATTTGCAAAATTTGGCAATTTGGAGGACCGAAAGGAACCAATGGCTTGTGT 1164  
Db 1144 CAGAGAGGGGTTGAGAAATCTTTTGAAGCTAGAGAACCGTAAAGAGCGGTGGCTGGTGGT 1203  
Qy 1165 GCTTTATGCACTGTTGCTGCTTCTGCAAGGCTATGGAAGGAATCTTATGTTGACTTTAGC 1224  
Db 1204 ACTTTAGCTCTTCTGCTGCTTCTGCAAGGCGATGGAAGCATGCTACGTCGAATTTGGC 1263  
Qy 1225 AGACAAGTTAGCAGGGTCAACAGGGATGAAGGTTGGAATAATTTAGAGCAGATCGAGAAC 1284

Db 1264 TGAGAAACTTTCGGG---AAAAGGAGTTAAAGTGGGAAATTCGAGCTGACGGTGAGCA 1320  
Qy 1285 GAAAGAAATTTGCAAGAGTGAACCTGCAATTTGGGAAGCTTCCCTACGATATTTATTTTCCC 1344  
Db 1321 GAAGGAGTTTGTCTAAGCAAGAGCTTCAGTTAGGAAGCTTCCGACGATACTTCTCTTCC 1380  
Qy 1345 AAAGCATTCCTCTCGGCCAACAAATAAGTATCCCTCAGAAAAGAGAGATGTTGATTCCTT 1404  
Db 1381 GAAAGAGCTCCACGG---GCTATTAAAGTACCCTTCAGAGCATAGAGATGTTGATTTCACT 1437  
Qy 1405 GATGGCAATTTGTAATCCCTTAAAGTCCCTTAAAGATGAG 1433  
Db 1438 CATGCTGTTGTGAATCTTCTTCGTTGAG 1466

RESULT 13  
AAC48356  
ID AAC48356 standard; DNA; 1685 BP.  
XX AAC48356;  
XX 18-OCT-2000 (first entry)  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 57179.  
DE Arabidopsis thaliana  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
XX EP1033405-A2.  
XX 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 20-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.

PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145591P.  
PR 27-JUL-1999; 99US-0145598P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 28-JUL-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.

PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 13-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match

35.5%; Score 578.4; DB 3; Length 1685;

Best Local Similarity 69.0%; Pred. No. 9e-156; Matches 887; Conservative 0; Mismatches 351; Indels 48; Gaps 5;	
QY	157 CAATTTATCTCAAGACGAGCTTGGTAAAGCCGTTAAAGCCGA---ACCTCCACGCAA 213
Db	219 CCATTGCTCAGAGAGGTTACTCTATGAAACCTTTAAACCGTGAATCACAATCAGCAG 278
QY	214 GGATTCATTTGCTCTCGCAGCAACAAACATCGTTGCTCTCTGAGA-----C 267
Db	279 CGAATCTTGGGTTACTCGTGTCTTCACTCTAAATGCTCTGAACTTGAAGGAGGAGG 338
QY	268 GAAAGAGGAGATTTTCAAGACAGATAGCCAGTGTCTCGACATGCTTCACTCTTGAAT 327
Db	339 AGAAGTTGAAGACTTTGAGCAACTTGTAAAGACCTTGAAGATGCTTCTCCACTGAAAT 398
QY	328 CATGATAGAGCCCTCGCAAAATTCGGCAACGACATAGCTATTGCTCTTCACTGTGCTGA 387
Db	399 CATGATTAAGCTCTTGAAGATTCGGAGACCAATCGCAATGCTTTTGTAGTGAGCTGA 458
QY	388 AGATGTTGCTTTGATGATGCAAAATTAAGCGGTCGACCCCTTTAGGGTTTTTCAGTTT 447
Db	459 AGATGTTGCTTTGATGATGCAAAATTAAGCGGTCGACCCCTTTAGGGTTTTTCAGTTT 518
QY	448 GGACACTGGGAGACTGAACCCAGAACTTATCACTTTTGTATGCGGTTGAGAGCAATTA 507
Db	519 AGATACAGGAGATTAACCCCTGAACCTGACAGCTCTTTGAGCGCAGTCGAGAGCAGTA 578
QY	508 TGAATTTGCGATTTGAGTACATGTTCCCTGATGCTTGTGAGTTTCAAGCTTTGTTGAGGAG 567
Db	579 CGGATTCGATTTGAGTACATGTTTCTGATGAGTTGAGTTCAAGCTTTTGTAGTGAGAA 638
QY	568 TAAGGGGTTATCTCTTTTACGAGGATGGGCCAACAGAGTGTTCAGGGTGAAGAGGT 627
Db	639 CAAGGGTTTCTCTCTATGATGATGCTTAAAGAGTGTCTCAAGAGTGTTCGCGTGTGAGGAA 698
QY	628 GAGGCTTTAAGAGGGCCCTTAAGGTTCTGAGGATGATTAATCTGCTGAGAGGAAAGA 687
Db	699 TAGACCTTTGCGTGTCTTAAAGGCTTAAAGCTTGAATTAACAGGACAGAGGAAAGA 758
QY	688 CCACTACCTGCTAGTGTCTGAAATPACCGGTTGTTTCAAGTGTGATCCGGCTTTTGAAGG 747
Db	759 CCAATCTCGGTTACGAGATCTGAGATCCCTATTTGTTCAAGTGTGATCCAGTGTGAGG 818
QY	748 AATGGATGTGGAATTTGGAAGCTTGGTGAAGTGAACCCCTGTTGCAAAATGTGAAGGCCA 807
Db	819 GTTAGATGGCGGTGTTGGAAGTCTTGTGAAGTGAATCTTTGCTTAATGTTGAAGGAGC 878
QY	808 TGACATATGGAATCTTCTTAGGACCATGAATGTGCTGTGAATTCCTTGCATGCAAAAGG 867
Db	879 TGATGTGTGAATCTTCTGAGAACTATGGAATGTTCCGGTGAATGCAATTTGCACACAAG 938
QY	868 ATATGTTTCCATTTGGGTGTGAGCCCTGCACTAGGCTGTTTACCTGGGCAACATGAAG 927
Db	939 GTATGTCTAATCGGTGTGAGCCGTGTACTAGCCGCTGCTTCCAGGCCCAATGAGAG 998
QY	928 GGAAGGAGGTGTGTGGGAGGATGCAAGAGTAAAGGATGTGCTTTCACAAAGGAAA 987
Db	999 AGAAGGAAGGTGTGTGGGAGATGCTAAAGCTTAAAGATGTGCTTACACAAAGGAA 1058
QY	988 TGTAAAGCAGCAGAAAGAGAGGAGTGTAAATGGAATGGCTATCCCAATCCCATGCAAA 1047
Db	1059 CATCAAGGAGGAGATGGTG-----CTGAGA 1085
QY	1048 TGGTGTATGCTTACCACGTGCTGACATTTTCAACAGCCCGAATGTAGTTAACTTGAAGCAG 1107
Db	1086 CTCAAAGCCCTGCTGTGCAAGAGATTTTGAAGCAACATGTGCTGCTTGTAGTGAAGC 1145
QY	1108 GACTGGAATTTGAAATTTGGCAAAATTTGGAGGACCGAAAGAACCATGGCTTCTGTGCT 1167
Db	1146 AGGAGGGTTGAGAACTTTTGAAGCTAGAGAACCGTAAAGAGCCGTGTTGTGCTACT 1205
QY	1168 TTATGACCAATGTGCTGCTTACGCAAGCTATGAGGAACTTATGTTGACTTACGAGA 1227

Db	1206 TTACGCTCTCTGGTGCCCTTTCTGCCAGGCGGAGGAGCATCGTACATCGAATTTGGCTGA 1265
QY	1228 CAAGTTAGCAGGGTCAACAGGGATGAAGTTGGAAATTTAGACGATCGAGAACAGAA 1287
Db	1266 GAACTTGGCGG---AAAGGAGTTAAAGTGGCGAAATTCGAGCTGACGGTGAAGCAGAA 1322
QY	1288 AGAATTTGCAAGAGTGAACCTGCAATTTGGGAAGCTTCCCTACGATATTTATTTTCCCAAA 1347
Db	1323 GGAGTTTGTCTAAGCAAGAGCTTCAGTTTGGGAGCTTCCCGCAGGATCTTCTCTTTCCGAA 1382
QY	1348 GCATTCGCTCTCGGCGCAACAATAAAGTATCCCTCAGAAAGAGAGATGTTGATTCCTTGAT 1407
Db	1383 AAGAGCTCCACGG---GCTATTAAAGTACCCCTTCAAGCATAGAGATGTTGATTCATCAT 1439
QY	1408 GGCATTTGTAATGCCTTAAAGATGAG 1433
Db	1440 GTCGTTTGTGNACTTCTTCGCTGAG 1465
RESULT 14	
ABZ14500	
ID	ABZ14500 standard; DNA; 1365 BP.
XX	ABZ14500;
AC	ABZ14500;
XX	21-JAN-2003 (first entry)
DT	XX
XX	XX
DE	Arabidopsis thaliana stress regulated gene SEQ ID NO 2305.
XX	XX
KW	Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX	XX
OS	Arabidopsis thaliana.
XX	XX
PN	W0200216655-A2.
XX	XX
PD	28-FEB-2002.
XX	XX
PF	24-AUG-2001; 2001WO-US026685.
XX	XX
PR	24-AUG-2000; 2000US-0227866P.
PR	26-JAN-2001; 2001US-0264647P.
PR	22-JUN-2001; 2001US-0300111P.
XX	XX
PA	(SCRI ) SCRIPPS RES INST.
PA	(SYGN ) SYNGENTA PARTICIPATIONS AG.
XX	XX
PI	Harper JF, Kreps J, Wang X, Zhu T;
XX	XX
DR	WPI; 2002-304127/34.
XX	XX
PT	Identifying a stress condition to which a plant cell has been exposed and
PT	producing plants with increased tolerance to these abiotic stresses.
XX	XX
PS	Claim 144; SEQ ID NO 2305; 577pp + Sequence Listing; English.
XX	XX
CC	The invention relates to identifying a stress condition to which a plant
CC	cell has been exposed, comprising: (a) contacting nucleic acid
CC	representative of expressed polynucleotides in the plant cell with an
CC	array or probes representative of the plant cell genome; and (b)
CC	detecting a profile of expressed polynucleotides in the plant cell
CC	characteristic of a stress response. The method is useful in the
CC	production of transgenic plants, cells and seeds and in producing plants
CC	with increased tolerance to abiotic stress. The present sequence is that
CC	of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC	in methods of the invention. Note: The sequence data for this patent is
CC	not represented in the printed specification but is based on sequence
CC	information supplied to Derwent by the European Patent Office
XX	XX
SQ	Sequence 1365 BP; 367 A; 240 C; 374 G; 384 T; 0 U; 0 Other;

Query Match 35.4%; Score 577.4; DB 6; Length 1365;  
Best Local Similarity 68.9%; Pred. No. 1.6e-155;  
Matches 886; Conservative 0; Mismatches 351; Indels 48; Gaps 5;

QY 157 CAATTTATCTCAAGACGGAGCTTGGTAAGCCGGTTAAACGCCGA---ACCTCCACGCAA 213  
Db 120 CCATTTGTCTCAGAGACGTTTACTCTATGAACCTTTTAAACCGCTGAGTCAATTCACGAAG 179  
QY 214 GGATTCATTTGCTCTCCGACGACCAACCATCGTTGCTTCTGCTTCTGAGA-----C 267  
Db 180 CGAATCTTGGGTACTCGTCTCACTCTAAATGCTCTCTGAAAGTTGAAGGAAGGAGG 239  
QY 268 GAAAGAGGAAGATTTTGAACAGATAGCCAGTGATCTCGACAATGCTTCACTCTTTGAAAT 327  
Db 240 AGAAGTTGAAGACTTTGAGCAACTTGTCTAAAGAGCTTGAAGATGCTTCTCCACTTGAAT 299  
QY 328 CATGATAGACCTTCACAAATTCGCAACGACATAGCTATTGCTTCACTGCTGCTG 387  
Db 300 CATGGATAAAGCTCTCAGAGATTCGAGACCAATCGCAATTCCTTTAGTGAGCTGA 359  
QY 388 AGATGTTGCTTTGATTCAGTATGCAAAATTCACGGGTCGACCCCTTTAGGGTTTTTCAGTTT 447  
Db 360 AGATGTTGCAATGATGANTATGACGTTTAACTGGAAGCCATTTAGGGTTTTTATGTTT 419  
QY 448 GGCACATGGGAGACTGAACCCAGAAACTTATCAACTTTTGTATGCGGTTGAGAAAGCATTA 507  
Db 420 AGATACAGGGAGATTAAACCTCGAAACGTAACGGCTCTTTGACGCACTCGAGAAGCAGTA 479  
QY 508 TGGAAATTCGATTCAGTACATGTTCCCTGATGCTGTTGAGGTTCCAGGCATTTGGTGAGAG 567  
Db 480 CGGAAATTCGAATTCAGTACATGTTTCTGATGCAAGTTTGAAGTTTCAAGCTTTTAGTGAGAA 539  
QY 568 TAAGGGGTTATCTCTTTCTACGAGGATGGCAACAGAGTGTTCAGGGGTGAGAAAGGT 627  
Db 540 CAAGGGTTTGTCTATCTATGAAGATGTCATCAAGAGTGTTCGCTGTGAGAAAGT 599  
QY 628 GAGCCCTTTAAGGAGGCGCTTAAGGCTCTCAGAGCATGGATAAATGCTCAGAGGAAGA 687  
Db 600 TAGACCTTTGCGTCTCTTAAGGGTCTTAAAGCTTGGATTACAGACAGAGGAAGA 659  
QY 688 CCACTACCTGGTACTAGTCTGAATACCGGTTGTTCAAGTTGATCCGGCTTTTGGGG 747  
Db 660 CCAATCTCCGGTACGAGATCTGAGATCCCTTATTTGTTCAAGTTGATCCAGTGTTTGAGG 719  
QY 748 AATGATGTTGGAATTTGGAAGCTTGGTGAAGTGAACCTCTGCAAAATGGAAGGCGCA 807  
Db 720 GTTAGATGCGCGTGTGGAAGCTTGTGGAAGTGAATCCCTTTGGCTAATGTTGAGGAGC 779  
QY 808 TGACATATGGAATCTCTTAGGACCATGAATGTGCTGTGTAATCTTTCATGCAATGCAAAAGG 867  
Db 780 TGATGTGTGGAATCTCTGAGAACTATGGAATGTTCCGGTGAATGCAATGCAACGACCAAGG 839  
QY 868 ATATGTTTCCATTTGGGTGTGAGCCCTGCACTAGGCTGTTTACCTGGGCAACATGAAG 927  
Db 840 GTATGTGTCAATCGGGTGTGAGCCGTGCTACTAGGCGGTGCTTCAGGCCCAACATGAGAG 899  
QY 928 GGAAGGAGGTGTGCTGGGAGGATGCCAAAGCTAAGGAATGTGCTCTTCAAAAGGAAA 987  
Db 900 AGAAGGAAGTGTGCTGGGAGATGCTTAAGCTTAAGAAATGTGCTTACACAAAGGAA 959  
QY 988 TGTAAGCAGCAGAAAGAGGAGGATGTTAATGGAATGGGCTATCCCAATCCCATGCAAA 1047  
Db 960 CATCAAGGAGGAAGTGTG-----CTGCAGA 986  
QY 1048 TGGTGATGCTACACTGTGCTGATCATTTTCAAGAGCCGATGTAGTTAACTTTGAGCAG 1107  
Db 987 CTCAAAGCCTGTGCTGTGCAAGAGATATTTGAAAGCAACAAATGTTGCTGATGAGCAA 1046  
QY 1108 GACTGGAATTCGAAATTTGGCAAAATTCGAGGACCGAAAGCAACATGCTGTGTGCT 1167  
Db 1047 AGGAGGGTTGAGAACTTTTGAAGCTTAGAGAACCGTAAAGGGCTGTGTTGCTGTA 1106  
QY 1168 TTATGACCATTGGTCCCTCTACTGCGAGGCTATGAGGAATCTTATGTTGACTTAGCAGA 1227  
Db 1107 TTAGCTCTTGGTCCCTTTCTGCGAGCGATGGAAGCATCGTACATCGAATTCGCTGA 1166

QY 1228 CAAATTAGCAGGCTCAACAGGATGAAGGTTGGAATAATTAGACAGATGAGAAACAGAA 1287  
Db 1167 GAACTTGGGG---AAAGGAGCTTAAAGTGGCGAAATTCGAGCTGACGCTGAGCAGAA 1223  
QY 1288 AGAATTTGCAAGAGTGAAGCTGCAATTTGGGAAGCTTCCCTACGATATATTTTCCCAA 1347  
Db 1224 GGAGTTTGTCTAAGCAAGAGCTTCAGTTAGGGAGCTTCCGACGATACTTCTCTTCCGAA 1283  
QY 1348 GCATTCGTCTCGGCCCAACATAAAGTATCCCTCAGAAACAGAGAGATGTTGATTCCTTGAT 1407  
Db 1284 AAGAGCTCCACGG---GCTATTAAAGTACCTTCAGACATAGAGATGTTGATTCACAT 1340  
QY 1408 GGCAATTTGTAATGCCCTTAAGATCA 1432  
Db 1341 GTCGTTGTGATCTTCTTCGGTGA 1365  
RESULT 15  
ADG88214  
ID ADG88214 standard; cDNA; 1365 BP.  
XX ADG88214;  
XX  
DT 22-APR-2004 (first entry)  
XX  
DE A. thaliana RPP4-upregulated pathogen infection-related gene #656.  
XX  
KW Pathogen infection-related gene; plant; Peronospora parasitica;  
KW defence mechanism; RPP4; pathogen resistance; transgenic plant; oomycete;  
KW fungus; bacterium; virus; nematode; insect; aphid; gene; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
FN WO200222675-A2.  
XX  
PD 21-MAR-2002.  
XX  
PF 14-SEP-2001; 2001WO-US028506.  
PR 15-SEP-2000; 2000US-0232778P.  
PR 22-JUN-2001; 2001US-0300183P.  
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.  
PA (UYNC-) UNIV NORTH CAROLINA.  
PA (GLAZ/) GLAZEBROOK J.  
PA (WANG/) WANG X.  
PA (DANG/) DANG J L.  
PA (EULG/) EULGEM T.  
PA (ZHUT/) ZHU T.  
XX Glazebrook J, Wang X, Dangl JL, Bulgem T, Zhu T;  
XX WPI; 2002-292409/33.  
DR  
XX Novel isolated polynucleotide, useful for conveying pathogen resistance  
XX to plants, and for identifying plants infected with a pathogen.  
PT  
XX Claim 3; SEQ ID NO 656; 605pp; English.  
PS  
XX The invention relates to 691 Arabidopsis thaliana genes (ADG87559--  
CC ADG87557)) whose expression is altered in response to pathogen infection,  
CC and to homologues of these genes from other plants or fungi, especially  
CC from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),  
CC cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The  
CC expression of genes of the invention was upregulated or downregulated in  
CC Arabidopsis plants infected with the oomycete Peronospora parasitica,  
CC indicating that they play a role in defence mechanisms. The genes of the  
CC invention are regulated by RPP7 or RPP8 which act via unconventional  
CC signalling cascades or by the RPP4-dependent pathway. The invention also  
CC relates to polypeptides encoded by the pathogen infection-related genes;  
CC promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)  
CC ; expression cassettes, host cells and pathogen-resistant transgenic  
CC plants and their progeny comprising a polynucleotide of the invention;



and a method of identifying a plant cell infected with a pathogen. The polynucleotide sequences and methods of the invention are useful for identifying plants infected with a pathogen, and for conferring resistance to pathogens such as oomycetes, fungi, bacteria, viruses, nematodes and insects (e.g., aphids). The present sequence represents an *Arabidopsis thaliana* gene whose expression is altered in response to *Peronospora parasitica* infection. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/pub/published\\_pct\\_sequences](http://wipo.int/pub/pub/published_pct_sequences).

SQ Sequence 1365 BP; 367 A; 240 C; 374 G; 384 T; 0 U; 0 Other;

Query Match 35.4%; Score 577.4; DB 6; Length 1365;

Best Local Similarity 68.9%; Pred. No. 1.6e-155;

Sequence	Matches	Mismatches	Indels	Gaps
Sequence 1	886	0	48	5
Sequence 2	886	0	48	5
Sequence 3	886	0	48	5
Sequence 4	886	0	48	5
Sequence 5	886	0	48	5
Sequence 6	886	0	48	5
Sequence 7	886	0	48	5
Sequence 8	886	0	48	5
Sequence 9	886	0	48	5
Sequence 10	886	0	48	5
Sequence 11	886	0	48	5
Sequence 12	886	0	48	5
Sequence 13	886	0	48	5
Sequence 14	886	0	48	5
Sequence 15	886	0	48	5
Sequence 16	886	0	48	5
Sequence 17	886	0	48	5
Sequence 18	886	0	48	5
Sequence 19	886	0	48	5
Sequence 20	886	0	48	5
Sequence 21	886	0	48	5
Sequence 22	886	0	48	5
Sequence 23	886	0	48	5
Sequence 24	886	0	48	5
Sequence 25	886	0	48	5
Sequence 26	886	0	48	5
Sequence 27	886	0	48	5
Sequence 28	886	0	48	5
Sequence 29	886	0	48	5
Sequence 30	886	0	48	5
Sequence 31	886	0	48	5
Sequence 32	886	0	48	5
Sequence 33	886	0	48	5
Sequence 34	886	0	48	5
Sequence 35	886	0	48	5
Sequence 36	886	0	48	5
Sequence 37	886	0	48	5
Sequence 38	886	0	48	5
Sequence 39	886	0	48	5
Sequence 40	886	0	48	5
Sequence 41	886	0	48	5
Sequence 42	886	0	48	5
Sequence 43	886	0	48	5
Sequence 44	886	0	48	5
Sequence 45	886	0	48	5
Sequence 46	886	0	48	5
Sequence 47	886	0	48	5
Sequence 48	886	0	48	5
Sequence 49	886	0	48	5
Sequence 50	886	0	48	5
Sequence 51	886	0	48	5
Sequence 52	886	0	48	5
Sequence 53	886	0	48	5
Sequence 54	886	0	48	5
Sequence 55	886	0	48	5
Sequence 56	886	0	48	5
Sequence 57	886	0	48	5
Sequence 58	886	0	48	5
Sequence 59	886	0	48	5
Sequence 60	886	0	48	5
Sequence 61	886	0	48	5
Sequence 62	886	0	48	5
Sequence 63	886	0	48	5
Sequence 64	886	0	48	5
Sequence 65	886	0	48	5
Sequence 66	886	0	48	5
Sequence 67	886	0	48	5
Sequence 68	886	0	48	5
Sequence 69	886	0	48	5
Sequence 70	886	0	48	5
Sequence 71	886	0	48	5
Sequence 72	886	0	48	5
Sequence 73	886	0	48	5
Sequence 74	886	0	48	5
Sequence 75	886	0	48	5
Sequence 76	886	0	48	5
Sequence 77	886	0	48	5
Sequence 78	886	0	48	5
Sequence 79	886	0	48	5
Sequence 80	886	0	48	5
Sequence 81	886	0	48	5
Sequence 82	886	0	48	5
Sequence 83	886	0	48	5
Sequence 84	886	0	48	5
Sequence 85				

QY 157 CAATTTATCTCAAAGACGGAGCTTGGTAAAGCCCGTTAACGCCGA---ACCTCCACGCAA 213

Db 120 CCATTGTCTCAGAGACGTTACTCTATGAAACCTTTAAACGCTGAGTCACATTCACGAAG 179

214 GGATTCCATTGTTCCCTCTCGCAGCAACAA CCA TCGTTGCTTCTGCTTCTTGAGA-----C 267

D<sub>b</sub>

180 CGAATCTTGGGTACCTCGTCCTTCACTGCTCCTCAGCTTGACACACACAC

2025

100

100

[illegible]

22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051

308 AGATGTTGCTTGATGGAGTATGCGAATATGACGGTCGACCTTTAGGGTTTCAGTTT 447

DD 360 AGA T G T T G C A T T G A T T G A A T A T G C A C G T T A A C T G G A A A G C C A T T A G G G T T T T A G T T T 419

448 GGACACTGGGAGACTGAACCCAGAACTTATCAACTTTTGTATGCGGTGAGAAGCATT 507  
QY

420 AGATACAGGGAGATTAAACCCCTGAAACGTCACAGGCTCTTTGACGCAGTCGAGAAAGCAGTA 479

QY 508 TGGAAATTCGCATTGAGTACATGTTCCCTGATGCTGTGAGGTTCAGGCATTGGTGAGGAG 567

Db 480 CGGGATTGGAATTGAGTACATGTTCTTGATGCAGTTGAGGTTCAAGCTTTAGTGAGGAA 539

QY 568 TAAGGGTTATTCTCTTTCTACGAGGATGGGCACCAAGAGTGTTCAGGGTGAGAAAGGT 627

Db 540 CAAGGGTTTGTCTCATTTCTATGAAGATGGTCAACAAGAGTGTCCGTGTGAGGAAAGT 599

628 GAGGCCCTTTAAGGAGGGCCCTTAAGGGTCTCAGAGCATGGATAACTGGTCAGAGGAAAGA 687

Db  
600 TAGACCTTTGCGTTCGTCTTAAAGGTCCTTAAAGCTTGGATTACAGGACAGCGAAAGA 659

[illegible]

660  
 661  
 662  
 663  
 664  
 665  
 666  
 667  
 668  
 669  
 670  
 671  
 672  
 673  
 674  
 675  
 676  
 677  
 678  
 679  
 680  
 681  
 682  
 683  
 684  
 685  
 686  
 687  
 688  
 689  
 690  
 691  
 692  
 693  
 694  
 695  
 696  
 697  
 698  
 699  
 700  
 701  
 702  
 703  
 704  
 705  
 706  
 707  
 708  
 709  
 710  
 711  
 712  
 713  
 714  
 715  
 716  
 717  
 718  
 719  
 720  
 721  
 722  
 723  
 724  
 725  
 726  
 727  
 728  
 729  
 730  
 731  
 732  
 733  
 734  
 735  
 736  
 737  
 738  
 739  
 740  
 741  
 742  
 743  
 744  
 745  
 746  
 747  
 748  
 749  
 750  
 751  
 752  
 753  
 754  
 755  
 756  
 757  
 758  
 759  
 760  
 761  
 762  
 763  
 764  
 765  
 766  
 767  
 768  
 769  
 770  
 771  
 772  
 773  
 774  
 775  
 776  
 777  
 778  
 779  
 780  
 781  
 782  
 783  
 784  
 785  
 786  
 787  
 788  
 789  
 790  
 791  
 792  
 793  
 794  
 795  
 796  
 797  
 798  
 799  
 800  
 801  
 802  
 803  
 804  
 805  
 806  
 807  
 808  
 809  
 810  
 811  
 812  
 813  
 814  
 815  
 816  
 817  
 818  
 819  
 820  
 821  
 822  
 823  
 824  
 825  
 826  
 827  
 828  
 829  
 830  
 831  
 832  
 833  
 834  
 835  
 836  
 837  
 838  
 839  
 840  
 841  
 842  
 843  
 844  
 845  
 846  
 847  
 848  
 849  
 850  
 851  
 852  
 853  
 854  
 855  
 856  
 857  
 858  
 859  
 860  
 861  
 862  
 863  
 864  
 865  
 866  
 867  
 868  
 869  
 870  
 871  
 872  
 873  
 874  
 875  
 876  
 877  
 878  
 879  
 880  
 881  
 882  
 883  
 884  
 885  
 886  
 887  
 888  
 889  
 890  
 891  
 892  
 893  
 894  
 895  
 896  
 897  
 898  
 899  
 900  
 901  
 902  
 903  
 904  
 905  
 906  
 907  
 908  
 909  
 910  
 911  
 912  
 913  
 914  
 915  
 916  
 917  
 918  
 919  
 920  
 921  
 922  
 923  
 924  
 925  
 926  
 927  
 928  
 929  
 930  
 931  
 932  
 933  
 934  
 935  
 936  
 937  
 938  
 939  
 940  
 941  
 942  
 943  
 944  
 945  
 946  
 947  
 948  
 949  
 950  
 951  
 952  
 953  
 954  
 955  
 956  
 957  
 958  
 959  
 960  
 961  
 962  
 963  
 964  
 965  
 966  
 967  
 968  
 969  
 970  
 971  
 972  
 973  
 974  
 975  
 976  
 977  
 978  
 979  
 980  
 981  
 982  
 983  
 984  
 985  
 986  
 987  
 988  
 989  
 990  
 991  
 992  
 993  
 994  
 995  
 996  
 997  
 998  
 999  
 1000  
 1001  
 1002  
 1003  
 1004  
 1005  
 1006  
 1007  
 1008  
 1009  
 1010  
 1011  
 1012  
 1013  
 1014  
 1015  
 1016  
 1017  
 1018  
 1019  
 1020  
 1021  
 1022  
 1023  
 1024  
 1025  
 1026  
 1027  
 1028  
 1029  
 1030  
 1031  
 1032  
 1033  
 1034  
 1035  
 1036  
 1037  
 1038  
 1039  
 1040  
 1041  
 1042  
 1043  
 1044  
 1045  
 1046  
 1047  
 1048  
 1049  
 1050  
 1051  
 1052  
 1053  
 1054  
 1055  
 1056  
 1057  
 1058  
 1059  
 1060  
 1061  
 1062  
 1063  
 1064  
 1065  
 1066  
 1067  
 1068  
 1069  
 1070  
 1071  
 1072  
 1073  
 1074  
 1075  
 1076  
 1077  
 1078  
 1079  
 1080  
 1081  
 1082  
 1083  
 1084  
 1085  
 1086  
 1087  
 1088  
 1089  
 1090  
 1091  
 1092  
 1093  
 1094  
 1095  
 1096  
 1097  
 1098  
 1099  
 1100  
 1101  
 1102  
 1103  
 1104  
 1105  
 1106  
 1107  
 1108  
 1109  
 1110  
 1111  
 1112  
 1113  
 1114  
 1115  
 1116  
 1117  
 1118  
 1119  
 1120  
 1121  
 1122  
 1123  
 1124  
 1125  
 1126  
 1127  
 1128  
 1129  
 1130  
 1131  
 1132  
 1133  
 1134  
 1135  
 1136  
 1137  
 1138  
 1139  
 1140  
 1141  
 1142  
 1143  
 1144  
 1145  
 1146  
 1147  
 1148  
 1149  
 1150  
 1151  
 115

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

100

1. The first step in the process is to identify the problem or issue that needs to be addressed. This involves gathering information and understanding the context of the problem.

100

DD TGA TGT GTG GAA CTT CTC GAG GAA CTA TGG ATG TCT CCG GTG AAT GCA TGT CAC GCA CAA GG 839  
 780

QY 868 ATATGTTCCATGGGTGGTGGACATGGCCGTGTTTACCTGGGCAACATGAAAG 927

D5 840 GTATGTGTCATCGGGTGTGAGCCGCTGCTTCCAGGCCAACATGAGAG 899

QY 928 GGAAGGGAGGTGGTGGTGGAGGATGCCAAAGCTAAGGAATGTGGTCTTCACAAAGGAAA 987

D**b** 900 AGAAGGAAGGTGGTGGGAAGATGCTAAAGCTAAAGAATGTGGTCTACACAAAGGGA 959

Qy	988	TGTTAAAGCAGCAGAAAGAGAGAGGATGTTAATGGAATGGGCTATCCCAATCCCATCGAAA	1047
Db	960	CATCAAGGGAGGAAGATGGT-----CTGCAGA	986
Qy	1048	TGGTGATGCTACCACTGTGCCTGCATCTTTCAACAGCGCCGAATGTAGTTAACTTGAGCAG	1107
Db	987	CTCAAGCCCTGCTGCTGTGCAAGGATATTGGAAGCACAATGTGGTTGCATTGAGCAA	1046
Qy	1108	GACTGGAAATTGAGAAATTGGCGAAAATTGGAGGACCGAAGAAACCAATGGCTTGTGTGCT	1167
Db	1047	AGGAGGGGTTGAGAAATCTTTTGAAGCTAGAGAAACCGTAAAGAGCGGTGGTGGTCTGCTACT	1106
Qy	1168	TTATGCACCATGGTGGCCCTACTGCCAGGCTATGGAGGAATCTTATGTTGACTTATGACGA	1227
Db	1107	TTACGCTCTCTTGTGGCCCTTTTGCGCAGGCGATGGAAGCATCGTACATCGAATTGGCTGA	1166
Qy	1228	CAAGTTAGCAGGGTCAACAGGGATGAAGSTTGGAAAAATTTAGAGCAGATGGAGAACAGAA	1287
Db	1167	GAAACTTTGCGG---AAAGAGGTTAAAGTGGCGAAATTCGAGCTGACGCTGAGCAGAA	1223
Qy	1288	AGAAATTTGCAAGAGTGAACTGCAAATTGGGAAGCTCCCTACGATATATTTTCCGAAA	1347
Db	1224	GGAGTTGCTTAAGCAAGAGCTTCAGTTAGGGAGCTTCCGACGATATCTCTCTTCCGAA	1283
Qy	1348	GCATTCGCTCTCGGCCAACAAATAAGTATCCCTCAGAAAAAGAGAGATGTTGATTCCTTGAT	1407
Db	1284	AAGAGCTCCAGG---GCTATTAAAGTACCCTTCAGAGCATGAGAGATGTTGATTCATCAT	1340
Qy	1408	GGCAATTTGTAATGCGCCTTAAGATGA	1432
Db	1341	GTCTGTTTGTGAATCTCTTCCGGTGA	1365

Search completed: February 20, 2006, 21:00:54  
Job time : 951 secs

**THIS PAGE LEFT BLANK**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 20:45:11 ; Search time 5977 Seconds  
(without alignments)

12751.574 Million cell updates/sec

Title: US-10-731-525-7

Perfect score: 1629

Sequence: 1 gcacggaggagaccacata.....taaaaaaaaaaaaaaaaaaa 1629

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_hic:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_est7:\*  
9: gb\_gss1:\*  
10: gb\_gss2:\*  
11: gb\_gss3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	600.8	36.9	765	2	BE658477 GM700006A
2	599.6	36.8	612	2	BE058404 snl5f07.y
3	570.6	35.0	1535	4	BN813913 Arabidops
4	567.2	34.8	1579	4	BN815501 Arabidops
5	542.2	33.3	608	2	BF070490 st20h11.y
6	539.4	33.1	552	2	BF071136 st12f06.y
7	510.2	31.3	516	3	B1316361 stf02c09
8	504	30.9	830	6	CB339330 CA23E1031
9	491.6	30.2	507	3	B1316488 saf03h04
10	486.6	29.9	959	7	CK264757 EST710835
11	481.4	29.6	629	1	AW508411 st40e11.y
12	474.8	29.1	704	6	CB340603 CA23E1011
13	473.4	29.1	910	6	CD573900 UCRPT01_0
14	450.8	27.7	764	6	CF119175 MTU10CS_P
15	450.8	27.7	764	6	CF119189 MTU10CS_P
16	446.2	27.4	968	7	CK263241 EST709319
17	438.4	26.9	716	7	CO492657 G.h.fbr-s
18	431.8	26.5	875	6	CB292552 UCRCS01_0
19	429.6	26.4	683	6	CA294320 MTU7CL_P4
20	429.2	26.3	660	6	CD487608 Gm ckr295
21	428.4	26.3	654	2	BG455000 NF112A09L
22	427	26.2	1299	10	CL972938 OsIFCC023

23	426.2	26.2	676	6	CA932068
24	426	26.2	858	2	BF253965
25	425	26.1	856	7	CK287268
26	420.2	25.8	836	9	BH652825
27	420.2	25.8	859	8	DN982440
28	415.6	25.5	963	7	CK286388
29	414.6	25.5	780	3	BM111117
30	413.6	25.4	689	7	CO105768
31	413.2	25.4	671	7	CO105768 GR_EB003
32	412.6	25.3	878	6	CF213409 CGF100062
33	412.4	25.3	729	6	CA808814
34	411.4	25.3	831	8	CK053400
35	411	25.2	608	8	CK172578
36	408.4	25.1	689	2	BG439877
37	408.2	25.1	606	6	CD713621
38	408.2	25.1	606	6	CD720052
39	407	25.0	653	2	BG124326
40	406.8	25.0	610	1	AJ558767
41	406.8	25.0	648	6	CD716959
42	406.8	25.0	769	2	BG581920
43	406.2	24.9	729	7	CV100088
44	404.2	24.8	756	10	DU052684
45	401.4	24.6	831	10	CW799574

## ALIGNMENTS

RESULT 1  
BE658477/c  
LOCUS  
DEFINITION  
mRNA sequence.  
765 bp mRNA linear EST 24-MAY-2001  
Gm-r1070 Glycine max cDNA clone Gm-r1070-2114 3',  
ACCESSION  
BE658477  
VERSION  
BE658477.1  
KEYWORDS  
EST.  
SOURCE  
Glycine max (soybean)  
ORGANISM  
Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
REFERENCE  
1 (bases 1 to 765)  
Vodkin,L., Kelm,P., Shoemaker,R., Retzel,E., Khanna,A., Coryell,V.,  
Erpelting,J., Raph,C., Shoop,E., Pardini,J., Liu,L. and Lewin,H.  
A Functional Genomics Program for Soybean (NSF 9872565)  
Unpublished (1999)  
Other ESTs: AW133461 corresponding to Gm-cl015-1660 (5')  
Contact: Vodkin, L.O., PI, A Functional Genomics Program for  
Soybean (NSF 9872565)  
Lewin, H. A., Director, Keck Center for Comparative and Functional  
Genomics  
University of Illinois  
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA  
Tel: (217) 244-6147  
Fax: (217) 333-4582  
Email: l-vodkin@uiuc.edu  
This clone is available through: Genome Systems, Inc. 4633 World  
Parkway Circle St. Louis, Missouri 63134. For further information  
call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)  
427-3324 or contact: clones@genomesystems.com or info@genome  
systems.com web site: www.genomesystems.com  
Seq primer: 5'-TTTTTTTTTTTTTTTTT(A/C/G)-3'.

## FEATURES

source  
1..765  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="Gm-r1070-2114"  
/clone\_lib="Gm-r1070"  
/notes="The library Gm-r1070 is a sequence-driven, reracked  
set of 9,216 clones selected from cDNA libraries from  
various tissues and stages of development of soybean that



```
Db 64 CCTGTGAATTCCTTGCATGCAAAAGGATATGTTTCCATTTGGTGTGAGCCCTGCACATAGG 123
Qy 902 CTTGTTTACCTGGGCAACATGAAAGGGAAGGAGGTGGTGGTGGAGGATGCCAAAGCT 961
Db 124 CTTGTTTACCTGGGCAACATGAAAGGGAAGGAGGTGGTGGTGGAGGATGCCAAAGCT 183
Qy 962 AAGGAATGTGGTCTTCAAAAGGAATGTAAAGCAGCAGAAAGAGGAGGATGTTAATGGA 1021
Db 184 AAGGAATGTGGTCTTCAAAAGGAATGTAAAGCAGCAGAAAGAGGAGGATGTTAATGGA 243
Qy 1022 AATGGGCTATCCCAATCCCAATGCAATGCTACCTACCTACCTACCTACCTACCTACCTACCT 1081
Db 244 AATGGGCTATCCCAATCCCAATGCAATGCTACCTACCTACCTACCTACCTACCTACCTACCT 303
Qy 1082 AGCCCGAATGTAGTTAATCTTGGCAGGACCTGGAATTCGAGAAATTTGGCAAAATTTGGAGGAC 1141
Db 304 AGCCCGAATGTAGTTAATCTTGGCAGGACCTGGAATTCGAGAAATTTGGCAGAAATTTGGAGGAC 363
Qy 1142 CGAAAGGAACATGGCTTGTGCTTTATGCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1201
Db 364 CGAAAGGAACATGGCTTGTGCTTTATGCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 423
Qy 1202 GAGGAATCTTATGTTAGCTTACGACAGCAAGTTAGCAGGCTCAACAGGATCAAGGATTTGGA 1261
Db 424 GAGGAATCTTATGTTAGCTTACGACAGCAAGTTAGCAGGCTCAACAGGATCAAGGATTTGGA 483
Qy 1262 AAATTTAGAGCAGATGAGAAACAGAAAGAAATTTGCAAAAGAGTGAATGCAATTTGGGAAGC 1321
Db 484 AAATTTAGAGCAGATGAGAAACAGAAAGAAATTTGCAAAAGAGTGAATGCAATTTGGGAAGC 543
Qy 1322 TTCCCTACGATATATTTTCCCAAGCATTCGCTCGGCCAACAAATAAAGTATCCCTCA 1381
Db 544 TTCCCTACGATATATTTTCCCAAGCATTCGCTCGGCCAACAAATAAAGTATCCCTCA 603
Qy 1382 GAAAAGA 1388
Db 604 GAAAAGA 610

RESULT 3
CNS0ADUO 1535 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSLTFB472C07 of Flowers and buds of strain col-0 of Arabidopsis
thaliana (thale cress).
ACCESSION BX813913
VERSION BX813913.1 GI:42474399
KEYWORDS HTC; GSUT_cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1535)
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schacher,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1535)
GENOSCOPE.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
```

```
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EP/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
FEATURES
Location/Qualifiers
source
1..1535
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSLTFB472C07"
/tissue_type="Flowers and buds"
/ecotype="Col-0"
/plasmid="pCMVSPORT 6"
complement(1..1535)
/gene="At1g62180"
```

```
ORIGIN
Query Match 35.0%; Score 570.6; DB 4; Length 1535;
Best Local Similarity 68.5%; Pred. No. 1.4e-146;
Matches 882; Conservative 0; Mismatches 359; Indels 46; Gaps 5;
Qy 157 CAATTTATCTCAAGAGCGGAGCTTGGTAAAGCCCGTTAAACGCCGA---ACCTCCACGCAA 213
Db 146 CCATTTGTCTCAGAGAGCTTACTCTATGAAACCTTTAAACGCTGAGTCACATTCCACGAAG 205
Qy 214 GGATTCATTTGCTCTCGCAGCAACCAACCATCGTTGCTTCTGCTTCTGAGA-----C 267
Db 206 CGAATCTTGGGTTACTCGTCTTCAACTCTAAATGCTCTGAAAGTTGAAGAGGAAGAGG 265
Qy 268 GAAAGAGGAGATTTTGAACAGATAGCAGTATCTGCAATGCTTCACTCTTCACTCTTGAAT 327
Db 266 AGAAGTTGAAGACTTTTGAGCACTTGTCTAAAGACTTGAAGATGCTTCTCACTTGAAT 325
Qy 328 CATGATAGAGCCCTCGACAAATTCGGCAACGACATAGCTATTGCTTCACTAGTGTGCTGA 387
Db 326 CATGGATTAAGCTCTTGGAGAGATTCGGAGACCAAAATCGCAATTTGCTTTAGTGAGCTGA 385
Qy 388 AGATGTTGCTTGAATGAGTATCGGAAATTTGAGGGTCGACCCCTTTAGGGTTTTCAGTTT 447
Db 386 AGATGTTGCAATGATTAATATGACGCTTTAACTGGAAAGCCATTTAGGGTTTTCAGTTT 445
Qy 448 GGACACTGGGAGACTGAACCCGAAACTTATCAACTTTTGTATGCGGTTTCAGAGGACATTA 507
Db 446 AGATCAGGAGAGATTAAACCTGAAACGCTACAGGCTCTTTGACGCGAGTCGAGAGCAAT 505
Qy 508 TGGAAATTCGCATTTAGTACATGTTCCCTGATGCTGTTGAGGTTTCAGGCATTTGGTGAGG 567
Db 506 CGGGATTTCGAATTTGAGTACATGTTTCTGATGAGTTGAGTTTCAAGCTTTAGTGAGGAA 565
Qy 568 TAAGGGGTTATTTCTTCTTACAGAGATGGGCAACAGAGTGTTCAGAGGTTGAGGAAAGGT 627
Db 566 CRAAGGGTTGTTTCTTCTTCTTATGAAGATGGTCTCAAGAGTGTTCGCGTGTGAGGAAAGT 625
Qy 628 GAGGCCCTTTAAGGAGGCGCTTAAAGGCTCTCAGAGCATGGAATGCTGTCAGAGGAAAGA 687
Db 626 TAGACCTTTGCGTGTGCTCTTAAAGGCTCTTAAAGCTTGAATTTACAGGACAGAGGAAAGA 685
Qy 688 CCAGTCACCTGTGCTAGGTCTGAAATATCCGGTTTGTTCAGGTTTGATCCCGCTTTTGAGGG 747
Db 686 CCAATCTCCGGGTACGAGATCTGAGATCCCTATTGTTTCAGTTGATCCAGTGTTTGAAGG 745
Qy 748 AATGGATGTGGAATTTGGAAGCTTGGTGAAGTGGAACCCCTGTTGCAAAATGTGAAGGCCA 807
Db 746 GTTAGATGGCGGTGTTGGAAGTCTTGTGAAGTGGAAATCTCTTTGCGCTAAATTTGAAGGAGC 805
Qy 808 TGACATATGGAATCTCTTTAGGACCATGAATGTCCTGTGAAATTCCTTGCATGCAAAAGG 867
Db 806 TGATGTGGAATCTTTCTGAGAACTATGGATGTTCCGGTGAATGCAATTCGCGCAAGG 865
```

```
Qy 868 ATATGTTTCATTTGGGTGTGAGCCCTGCATAGSCCTGTTTACCTGGGCAACATGAAG 927
Db 866 GTATGTGTCAATCGGCTGTGAGCCGTGTACTAGCCCGTGTCTCCAGGCCCAACATGAG 925
Qy 928 GGAAGGAGGTGGTGGTGGGAGGATGCAAAAGCTTAAGCAATGTGCTTCCACAAAGGAAA 987
Db 926 AGAAGGAAGTGGTGGTGGGAGATGCTAAAGCTTAAGCAATGTGCTTCCACAAAGGAAA 985
Qy 988 TGTAAACAGCAGAAAGAGGAGGATGTTAATGGAAATGGGCTATCCCAATCCCATGCAAA 1047
Db 986 CATCAAGGAGGAGATGGTG-----CTGCAGA 1012
Qy 1048 TGGTGATGCTACCACTGTGCTGACATTTTCAACAGCCCCGAAATGAGTAACTTGAGCAG 1107
Db 1013 CTCAAAGCCTGCTGCTGTGCAAGAGATATTGAAAGCAACAATGTGTTGCAATGAGCAA 1072
Qy 1108 GACTGGAAATTTGAGAAATTTGGCAAAATTTGGAGGACCCGAAAGCAACATGGCTTGTGTGCT 1167
Db 1073 AGGAGGGGTGAGAAATCTTTTGAAGCTAGAGAACCGTGAAGAGCGGTGTTGTGCTGCTACT 1132
Qy 1168 TTATGCACCATGTTGCCCTTACTCCAGGCTATGAGGAAATCTTATGTTGACTTTAGCAGA 1227
Db 1133 TTACGCTCTTGTGTGCTTCTGCAAGGCTGGAAGCATCGTACATCGAAATTTGGCTGA 1192
Qy 1228 CAAGTTAGCAGGTCACACGGATGAAGTTTGGAAATTTAGAGCAGATGAGAGAACAGAA 1287
Db 1193 GAAACTTTCGGG-----AAAGAGAGTTAAAGTGGCGAAATTCGAGCTGACGCTGAGCAGAA 1249
Qy 1288 AGAATTTTGAAGAGTGAACCTGCAATTTGGGAAGCTTCCCTACGATA-TTATTTTTCCAA 1346
Db 1250 GGAGTTTGCTNAGCAAGAGCTTCAGTTAGGAGCTTCCCAAGATCTTCTCTTTCGG 1309
Qy 1347 AGCATTCGCTCGGCCCAACAAATAAAGTATCCCTCAGAAAAGAGAGATGTTGATTCCTGA 1406
Db 1310 AAAAGAGCTGGCAGCGGCTATTAAAGTACCCTTCAGAGCATAGATATGCTGATTCACCTCA 1369
Qy 1407 TGGCATTTGTAATGCTTAAAGTGAAG 1433
Db 1370 TGTGTTTGTGAATCGTCTTCGGTGAG 1396

RESULT 4
CNSOACVY 1579 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA complete sequence from clone
DEFINITION GSLTSL682B02 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).
BX815501
BX815501.1 GI:42473038
HTC; GSLT_cDNA.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1579)
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished
REFERENCE 2 (bases 1 to 1579)
Genoscope.
Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
```

```

V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
FEATURES
Location/Qualifiers
1..1579
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSLTSL682B02"
/tissue_type="Adult vegetative tissue"
/ecotype="Col-0"
/plasmid="pCMVSPORT 6"
complement(1..1579)
/gene="Atlg62180"

gene
ORIGIN
Query Match 34.8%; Score 567.2; DB 4; Length 1579;
Best Local Similarity 68.4%; Pred. No. 1.3e-145;
Matches 880; Conservative 0; Mismatches 358; Indels 48; Gaps 5;

Qy 157 CAATTATCTCAAGACGGAGCTTGGTAAAGCCGTTAAACGCCGA---ACCTCCACCA 213
Db 209 CCATTTGTCTCAGACAGCTTACTCTATGAGACCTTTAAACGCTGAGTCACATTCACGAAG 268
Qy 214 GGATTTCCATTTCTCTCCAGCAACAACATCGTTGCTTCTGCTTCTGAGA-----C 267
Db 269 CGAATCTGGGTACTCGTCTTCACTCTAATTTGCTCTCTGAAGTGAAGAAAGGAGG 328
Qy 268 GAAAGAGGAAGATTTTGAACAGATAGCAGTGTATCTCGACAATGCTTCACTCTTGGAAT 327
Db 329 AGAATTTGAAGACTTTTGAGCAACTTGTCTAAGGAGCTTGAAGATGCTTCTCCACTTGAGAT 388
Qy 328 CATGGATAGAGCCCTCGACAAATTCGGCAAGCAGATAGCTATTGCTTCACTGCTGCTGA 387
Db 389 CATGGATAAGGCTCTTGAGAGATTCGGAGACCAAAATCGCAATTCCTTTAGTGAGCTGA 448
Qy 388 AGATTTGCTTTGATTGAGTATGCAAAATTCAGCGGTTCGACCCCTTTAGGGTCTTTCACTTT 447
Db 449 AGATTTGCTTTGATTGAGTATGCAAAATTCAGCGGTTCGACCCCTTTAGGGTCTTTCACTTT 508
Qy 448 GGACACTGGGAGACTGAACCCAGAAACTTATCAAACTTTTGTATGCGGTTCGAGAGCAATTA 507
Db 509 AGATACAGGGAGATTAACCCCTGAGACGTACAGGCTCTTTGACGCAAGTCGAGAAAGCAGTA 568
Qy 508 TGGAAATTCGCAATTCAGTACATGTTCCCTGATGCTGTTGAGGTTCCAGGATTCAGGAGGAG 567
Db 569 CGGATTTGCAATTCAGTACATGTTTCTGATGCTGAGTTTCAAGCTTTAGTGAGGAA 628
Qy 568 TAAGGGGTATTCTCTTTTACGAGGATGGGCACCAAGAGTGTTCGAGGGTTCGAGAAAGGT 627
Db 629 CAAGGGTTGTTCTCTCTATCTATGAAGTGTGTCATCAAGAGTGTTCGCTGTGAGGAAAGT 688
Qy 628 GAGGCCCTTTAGGAGGGCCCTTAAAGGTTCTCAGAGCATGGAATACTGGTTCAGAGGAAAGA 687
Db 689 TAGACCTTTGCGTCTCTTTGAGGGTCTTAAAGCTTGGATTGAGTACAGACAGAGGAAAGA 748
Qy 688 CCAGTCACTGGTACTAGGTTGAAATACCGGTTGTTTCAGGTTGATCCGGCTTTTGAGGG 747
Db 749 CCAATCTCCGGGTACGAGATCTGAGATCCCTTATTTGTCAGGTTGATTCAGTGTTTGAGAGG 808
Qy 748 AATGATGCTGGAATTTGGAAGCTTTGGTGAAGTGTGCAACCTCTTGTCAAAATGTGAAGGCCA 807
Db 809 GTTAGATGGCGGTGTTTGAAGTCTTGTGAAGTGGAAATCTTTGGCTAATGTTGAAGGAGC 868
Qy 808 TGACATATGGAATCTTCCTTAGGACCAATGAATGTCGTGTAATTCCTTGGCATGCAAAAGG 867
Db 869 TGATGTGTGGAACCTTTCTGAGAACTACGAGATGTTCCGGTGAATGCAATGCAATGCACGCAAGG 928
```



QY 868 ATATGTTTCCATTTGGGTGAGCCCTGCACTAGGCGCTGTTTACCTGGCAACATGAAG 927  
 DB 929 GTATGTGTCATCGGTGTGAGCGGTGTACTAGGCGGTGCTTCCAGGCCCAACATGAG 988  
 QY 928 GGAAGGAGGTGTGTGGGAGGATGCAAGCTTAAGCATGTGCTTCCAAAGAAA 987  
 DB 989 AGAAGGAAGGTGTGTGGGAGGATGCTTAAGCTTAAGCATGTGCTTCCAAAGGAA 1048  
 QY 988 TGTAAAGCAGCAGAAAGAGGAGGATGTTAATGGAATGGCTATCCCAATCCCATGCAA 1047  
 DB 1049 CATCAAGGAGGAGATGGTG-----CTGCAGA 1075  
 QY 1048 TGTGATGTACCACTGTGCTGACATTTTCAACAGCCCGAATGTAGTTAACTTGAGCAG 1107  
 DB 1076 CTCAAAGCCTGTGCTGTGCAAGAGATATTTGAAAGCAACAATGTGTTGCAATTGAGCAA 1135  
 QY 1108 GACTGGAATTGAGAAATTTGCAAAATTTGGAGGACCGGAAGAACCATGGCTGTGTGCT 1167  
 DB 1136 AGAGGGGTGTGAGAAATCTTTGGAAGCTAGAGAACCGTAAAGAGCGGTGTGCTGCTACT 1195  
 QY 1168 TTATGACACCATGTGTCCTTACTGCCAGGCTATGGAGGAATCTTATGTTGACTTAGCAGA 1227  
 DB 1196 TTACGCTCTTGTGTGCTTCTGCCAGGCGATGGAAGCATGTACATCGAATGTGCTGA 1255  
 QY 1228 CAAGTTAGCAGGTCACAGGATGAAGTTGGAATTTTAGCAGCATGGAGAACAGAA 1287  
 DB 1256 GAAACTTGCAGG---AAAAGGAGTTAAAGTGGCGMAATTCGAGCTGACGTCAGCAGAA 1312  
 QY 1288 AGAATTTGCAAGAGTGAATGCAATTTGGAAGCTTCCCTACGATATATTTTCCAAA 1347  
 DB 1313 GGAGTTTGTGTAAGCAAGAGCTTCAGTTAGGAGCTTCCGACGATATCTTATTTCCGAA 1372  
 QY 1348 GCATTGCTCTCGGCCAACATAAAGTATCCTCAGAAAGAGAGATGTTGATTCCTTGAT 1407  
 DB 1373 AAGAGCTCCAGG---GCTATTAGTACCTTCAGAGCATAGAGATGTTGATTCACATCAT 1429  
 QY 1408 GCATTGTTAAATGCCCTTAAGATGAG 1433  
 DB 1430 GTGCTTGTGAATCTTCTTCGCTGAG 1455

## RESULT 5

BF070490

LOCUS

DEFINITION

BF070490 608 bp mRNA linear EST 13-JUL-2004  
 st20h11.v1 Gm-cl065 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
 Gm-cl065-1918 5' similar to TR:Q9ZP22 Q9ZP22 APS REDUCTASE  
 PRECURSOR. ;, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Glycine max (soybean)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 rosids; eurosids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.

1 (bases 1 to 608)

Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V.,

Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,  
 Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,  
 Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,  
 Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,  
 McCann, R., Waterston, R. and Wilson, R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com) Insert length: 3116 Std Error: 0.00 High quality sequence stop: 386.

## FEATURES

source

1..608  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /cultivar="Williams"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl065-1918"  
 /tissue\_type="germinating shoots"  
 /lab\_host="DH10B"  
 /clone\_lib="Gm-cl065"  
 /note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from germinating shoots of the cultivar Williams. The seeds were allowed to germinate for 24 hours prior to being cold stressed for 2 days at 4C. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

## ORIGIN

Query Match 33.3%; Score 542.2; DB 2; Length 608;  
 Best Local Similarity 94.3%; Pred. No. 8.5e-139;  
 Matches 562; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 8 GAGAGAACCCATAACAGCTAGTTAATGGCCCTCGCTTCACTTCTTCAATTCGGCACCA 67  
 DB 12 GAGAGAACCCATAACAGCTAGTTAATGGCCCTCGCTTCACTTCTTCAATTCGGCACCA 71  
 QY 68 ACTTCCACCTTCCCATCATCGGAACCCAAATTCGCAAAATTCGGTCAATTAGATTTCG 127  
 DB 72 ACTTCCACCTTCCCATCATCGGAACCCAAATTCGCAAAATTCGGTCAATTAGATTTCG 131  
 QY 128 GAGAGCCCATATGGAGCGCGCTTAATTCAAATTTATCTCAAAGACGAGCTTGGTAAAG 187  
 DB 132 GAGAGCCCATATGGAGCGCGCGCTTAATTCAAATTTATCTCAAAGACGAGCTTGGTAAAG 191  
 QY 188 CCGTTTAAACGCGGAACCTCCAGCAAGGATTCATTTGCTCTCGCAGCAACCAATC 247  
 DB 192 CCGTTTAAACGCGGAACCTCCAGCAAGGATTCATTTGCTCTCGCAGCAACCAATC 251  
 QY 248 GTTGCTTCTGCTCTCTAGAGGAGGAGGATTTTGAACAGATAGCCAGTGTCTCGAC 307  
 DB 252 GTTGCTTCTGCTCTCTAGAGGAGGAGGATTTTGAACAGATAGCCAGTGTCTCGAC 311  
 QY 308 AATGCTTCACTCTTCAAAATCATGGATAGAGCCCTCGACAAATTCGCAACGACATAGCT 367  
 DB 312 AATGCTTCACTCTTCAAAATCATGGATAGAGCCCTCGACAAATTCGCAACGACATAGCT 371  
 QY 368 ATTGCTTTCAGTGGTGTCTGAAGATGTTGCTTTGATGAGTATCGGAATTCGAGGGTGA 427  
 DB 372 ATTGCTTTCAGTGGTGTCTGAAGATGTTGCTTTGATGAGTATCGGANATTCGAGGGTGA 431  
 QY 428 CCTTTAGGGTTTTCAGTTTGGACACTGGGAGACTGAACCCAGAAATTCATCAACTTTT 487  
 DB 432 CCTTTAGGGTTTTCAGTTTGGACACTGGGAGACTGAACCCAGAAATTCATCAACTTTT 491  
 QY 488 GATGCGGTTGAGAAGCATTATGGAATTCGATTGAGTACATGTTCCCTGATGCTGTTGAG 547  
 DB 492 GATGCGGCTGAGAAGCATTATGGAATTCGATTGAGTACATGTTCCCTGATGCTGTTGTC 551  
 QY 548 GTTCAGGCATTGCTGAGGAGTAGAGGGTTATTCTCTTTCTACGAGGATGGGCACCA 603

```

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
552 GTTCAAGCATGTGTGAGGAGTAAAGGTATCTCTTTTACCAAGGATGGCTCCCA 607

Db
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
552 GTTCAAGCATGTGTGAGGAGTAAAGGTATCTCTTTTACCAAGGATGGCTCCCA 607

RESULT 6
BF071136
LOCUS
DEFINITION
BF071136 552 bp mRNA linear EST 13-JUL-2004
Gm-cl065-y1 Gm-cl065 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-cl065-1116 5' similar to TR:048887 048887 5'-ADENYL SULFATE
REDUCTASE. [1], mRNA sequence.
ACCESSION
BF071136
VERSION
BF071136.1 GI:10845541
KEYWORDS
EST.
SOURCE
Glycine max (soybean)
ORGANISM
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE
1 (bases 1 to 552)
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V.,
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. This clone is
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Insert Length: 1292 Std Error: 0.00
High quality sequence stop: 417.
FEATURES
source
location/Qualifiers
1..552
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl065-1116"
/tissue_type="germinating shoots"
/lab_host="DH10B"
/clone_lib="Gm-cl065"
/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from mRNA isolated
germinating shoots of the cultivar Williams. The seeds
were allowed to germinate for 24 hours prior to being
cold stressed for 2 days at 4C. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with a XhoI restriction site. EcoRI
adapters were ligated to the blunt-ended cDNA fragments
followed by XhoI digestion. The cDNA fragments were
directionally cloned into the EcoRI-XhoI restriction site
of the pBluescript vector. The ligated cDNA fragments were
transformed into DH10B host cells (GibcoBRL). This library
was constructed in the laboratory of Dr. Randy
Shoemaker."
ORIGIN
Query Match 33.1%; Score 539.4; DB 2; Length 552;
Best Local Similarity 98.5%; Pred. NO. 4.9e-138;
Matches 543; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 311 GCTTACCTCTTGAATCATGATAGAGCCCTCGACAAATTCGGCAACGACATAGCTATT 370
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

```

```

Db
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1 GCTTACCTCTTGAATCATGATAGAGCCCTCGACAAATTTCGGCAACGACATAGCTATT 60

Qy
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
371 GCCTTCAGTGTGCTCTGAAGATGTTGCTTTGATGATGCGAAATTCGCGGTTCGACCC 430

Db
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
61 GCCTTCAGTGTGCTCTGAAGATGTTGCTTTGATGATGCGAAATTCGCGGTTCGACCC 120

Qy
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
431 TTTTAGGGTTTTTCAGTTTGGACACTGGGAGACTGAACCCGAAACTTATCAACTTTTGTAT 490

Db
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
121 TTTTAGGGTTTTTCAGTTTGGACACTGGGAGACTGAACCCGAAACTTATCAACTTTTGTAT 180

Qy
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
491 GCGGTTGAGAAGCATTTATGGAATTCGCATTGAGTACATGTTCCCTGATGCTGTGAGGTT 550

Db
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
181 GCGGTTGAGAAGCATTTATGGAATTCGCATTGAGTACATGTTCCCTGATGCTGTGAGGTT 240

Qy
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
551 CAGGCATTGTGAGGAGTAAGGGTTATTCTCTTTCTACGAGGATGGCCACCAAGATGT 610

Db
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
241 CAGGCATTGTGAGGAGTAAGGGTTATTCTCTTTCTACGAGGATGGCCACCAAGATGT 300

Qy
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
611 TGCAGGCTGAGAAAGGTGAGGCTTTTAAAGGAGGCGCCCTTAAGGGTCTCAGAGCATGGATA 670

Db
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
301 TGCAGGCTGAGAAAGGTGAGGCTTTTAAAGGAGGCGCCCTTAAGGGTCTCAGAGCATGGATA 360

Qy
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
671 ACTGGTCAGAGGAAAGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 730

Db
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
361 ACTGGTCAGAGGAAAGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

Qy
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
731 GATCCGCTTTTTCAGGGAATGGATGGTGAATTTGGAAGCTTGGTGAATGGTGAATGGTGAAT 790

Db
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
421 GATCCGCTTTTTCAGGGAATGGATGGTGAATTTGGAAGCTTGGTGAATGGTGAATGGTGAAT 480

Qy
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
791 GCAAACTGAGAGGCCCATGACATATGAACTTCCTTAGGACCATGAATGTGCTGTGAAT 850

Db
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
481 GCANATGTAGCGGCCCATGACATATGAACTTCCTTAGGACCATGAATGTGCTGTGAAT 540

Qy
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
851 TCCTTGCATGC 861

Db
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
541 TCCTTGCATGC 551

RESULT 7
BF1316361
LOCUS
DEFINITION
BF1316361 516 bp mRNA linear EST 21-JUL-2004
saf02c09.y1 Gm-cl065 Glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-cl065-4337 5' similar to TR:Q39248 Q39248
3'-PHOSPHODENOSINE 5'-PHOSPHOSULFATE REDUCTASE ;, mRNA sequence.
ACCESSION
BF1316361
VERSION
BF1316361.1 GI:14990688
KEYWORDS
EST.
SOURCE
Glycine max (soybean)
ORGANISM
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE
1 (bases 1 to 516)
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V.,
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. This clone is
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Insert Length: 1292 Std Error: 0.00
High quality sequence stop: 417.
FEATURES
source
location/Qualifiers
1..552
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl065-1116"
/tissue_type="germinating shoots"
/lab_host="DH10B"
/clone_lib="Gm-cl065"
/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from mRNA isolated
germinating shoots of the cultivar Williams. The seeds
were allowed to germinate for 24 hours prior to being
cold stressed for 2 days at 4C. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with a XhoI restriction site. EcoRI
adapters were ligated to the blunt-ended cDNA fragments
followed by XhoI digestion. The cDNA fragments were
directionally cloned into the EcoRI-XhoI restriction site
of the pBluescript vector. The ligated cDNA fragments were
transformed into DH10B host cells (GibcoBRL). This library
was constructed in the laboratory of Dr. Randy
Shoemaker."
ORIGIN
Query Match 33.1%; Score 539.4; DB 2; Length 552;
Best Local Similarity 98.5%; Pred. NO. 4.9e-138;
Matches 543; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 311 GCTTACCTCTTGAATCATGATAGAGCCCTCGACAAATTCGGCAACGACATAGCTATT 370
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

```

57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)  
Seq primer: -40RP from Gibco  
High quality sequence stop: 396.

## FEATURES

Location/Qualifiers  
1..516  
/organism="Glycine max"  
/mol\_type="mRNA"  
/cultivar="Williams"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-c1065-4337"  
/tissue\_type="germinating shoots"  
/lab\_host="DH10B"

/clone\_lib="Gm-c1065"  
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from germinating shoots of the cultivar Williams. The seeds were allowed to germinate for 24 hours prior to being cold stressed for 2 days at 4C. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

## ORIGIN

Query Match 31.3%; Score 510.2; DB 3; Length 516;  
Best Local Similarity 99.4%; Pred. No. 6.2e-130;  
Matches 512; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 21 ACAGCTAGTTAATGGCCCTCGCTTTCACCTTCTCAATTCGCGACCAACTTCCACCTTCC 80
DB 2 ACAGCTAGTTAATGGCCCTCGCTTTCACCTTCTCAATTCGCGACCAACTTCCACCTTCC 61
QY 81 CATCATCGGAACCAACTTCCGCAATTTGGGTCAATTAGGATTTGGAGAGGCCCATTTG 140
DB 62 CATCATCGGAACCAACTTCCGCAATTTGGGTCAATTAGGATTTGGAGAGGCCCATTTG 121
QY 141 GAGCGCGGTAATTTCAATTTATCTCAAGAGCGAGCTGGTAAAGCCGTTAAAGCCCG 200
DB 122 GAGCGCGGTAATTTCAATTTATCTCAAGAGCGAGCTGGTAAAGCCGTTAAAGCCCG 181
QY 201 AACCTCCAGCAAGGATTCATTTGCTTCTCCGCAACCAACCATCGTTGCTTCTGCTT 260
DB 182 AACCTCCAGCAAGGATTCATTTGCTTCTCCGCAACCAACCATCGTTGCTTCTGCTT 241
QY 261 CTGAGACGAAAGGAGGATTTTGAACAGATAGCCAGTATCTCGACAATGCTTCAACCTC 320
DB 242 CTGAGACGAAAGGATTTTGAACAGATAGCCAGTATCTCGACAATGCTTCAACCTC 301
QY 321 TTGAATATCATGTAGAGCCCTCGACAAATTCGGCAACGACATAGCTATTCAGTG 380
DB 302 TTGAATATCATGTAGAGCCCTCGACAAATTCGGCAACGACATATCTATTCGCTTCA 361
QY 381 GTGCTGAAGATGTGCTTTGATTTGAGTATCGAAATTTGACGGGTGACCCCTTTAGG 440
DB 362 GTGCTGAAGATGTGCTTTGATTTGAGTATCGAAATTTGACGGGTGACCCCTTTAGG 421
QY 441 TCAGTTTGGACACTGGGAGACTGAAACCCAGAAACTTATCAACTTTTGTATCGGTTGAG 500
DB 422 TCAGTTTGGACACTGGGAGACTGAAACCCAGAAACTTATCAACTTTTGTATCGGTTGAG 481
QY 501 AGCATTTAGAAATTCGCATTTAGTATGATGATGATGATGATGATGATGATGATGATG 535
DB 482 AGCATTTAGAAATTCGCATTTAGTATGATGATGATGATGATGATGATGATGATGATG 516
```

## RESULT 8

CB339330

LOCUS

830 bp mRNA linear EST 14-MAR-2003

## DEFINITION

ACCESSION CB339330  
VERSION CB339330.1  
KEYWORDS GI:28958671  
SOURCE Vitis vinifera  
ORGANISM Vitis vinifera

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

CA23EI03IVFb B07 Cabernet Sauvignon Leaf - CA23EI Vitis vinifera  
cDNA clone CA23EI03IVFb\_B07 5', mRNA sequence.

CB339330  
CB339330.1 GI:28958671

EST.  
Vitis vinifera

Vitis vinifera

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; Vitaceae; Vitis.

1 (bases 1 to 830)

Goes da Silva, F., Lim, H., Iandolino, A., Baek, J., Leslie, A., Xu, J.,  
Jones, K., Walker, M. A. and Cook, D. R.

Transcriptional responses of Vitis vinifera to infection by the  
bacterial pathogen Xylella fastidiosa

Unpublished (2003)

Contact: Douglas Cook, PhD

CABES Genome Facility

UC Davis, Plant Pathology

One Shields Ave, Davis, CA 95616, USA

Tel: 530 754 6561

Fax: 530 754 6617

Email: drcook@ucdavis.edu

Seq primer: ACGTACCGGACATATGCC.

Location/Qualifiers

1..830

/organism="Vitis vinifera"

/mol\_type="mRNA"

/cultivar="Cabernet Sauvignon"

/db\_xref="taxon:29760"

/clone="CA23EI03IVFb\_B07"

/sex="Hexamphrodite"

/dev\_stage="Mid-season leaf material, collected July 25,  
2001."

/lab\_host="DH5alpha"

/clone\_lib="Cabernet Sauvignon Leaf - CA23EI"

/notes="Organ: Leaf; Vector: pDNR; Site 1: SfiI; Site 2:  
SfiI; CA23EI is a cDNA library of Cabernet Sauvignon

leaves. The leaves were collected on July 25, 2001, in

Napa Valley, California, and represent leaves in

mid-season development. These leaves were verified to be

infected with the bacterial pathogen, Xylella fastidiosa,

based on a diagnostic assay using PCR and Xylella-specific

primer pairs. The plants were asymptomatic at the time of

collection, but later developed symptoms. cDNAs were made

by oligo-dT priming and directionally cloned. 5' and 3'

adaptors were used in cloning as follows:

5'-AACGATGGTATCACGAGTGGCCATTCACGCCGGG-3' and

5'-ATTCTAGAGCCGAGCGCGGACATG-dt(30)NN-3'. Library was

constructed using the Clontech Creator SMART kit and

size-selected to contain the 0.5-3 kb size fraction."

## ORIGIN

Query Match 30.9%; Score 504; DB 6; Length 830;  
Best Local Similarity 76.7%; Pred. No. 3.7e-128;  
Matches 647; Conservative 0; Mismatches 179; Indels 18; Gaps 2;

```
QY 270 AAGAGGAAGATTTTGAACAGATAGCCAGTATCGACATGCTTCCCTTTGAAATCA 329
DB 4 AGGTGGAGGACTATGAGCAATTTGGCAATGACATTTGCTTTTCAATTA 63
QY 330 TGGATAGAGCCCTCGCAAAATTTCCGCAACGACATAGCTATTGCTTTCAGTGGTCTGAAG 389
DB 64 TGGATAGAGCCCTTTGGCAATTTGGCAATGACATTTGCTTTTCAATTA 123
QY 390 ATGTTGCTTTGATTTGAGTATGCGAAATTTGACGGGTGACCCCTTTAGGTTTTCAGTTGG 449
DB 124 ATATTGCTTTGATTTGAAATATGACGCTTAACCTGCTCGTTTATGGGTATTGAGCTGG 183
QY 450 ACATCTGGGAGACTGAACCCAGAAACTTATCAACTTTTGTATGCGGTTGAGAAGCATTTATG 509
DB 184 ACATCTGGAGGTTGAACCCGGAACATATCAATTTCTTTTGACCGGTTTGAGAAACATTTATG 243
```

```
QY 510 GAATTCGATTGAGTACATGTTCCCTGATGCTGTGAGGTTTCAGGCATTGGTGAGGAGTA 569
Db 244 GCATCCGCAATTGATACATGTTTCCAGATGCTGTGAGTTTCAGGCTTGGTGAGGAGCA 303
QY 570 AGGGGTTTATTTCTCTTTCTACGAGGATGGGACCAAGAGTGTGTCAGGTTGAGAAAGTGA 629
Db 304 AGGAGCTGTTTCTCTCTCTACGAGGATGGGACCAAGAGTGTGTCAGGTTGAGAAAGTGA 363
QY 630 GGCCTTTAAGAGGGCCCTTAAGGCTCTCAGAGCATGATGTAAGTGTGTCAGAGGAAAGACC 689
Db 364 GACCTTTGAGAGGGCACTGAAGGTTTGGCTGCTGGATCAGAGGCAAGAGAGATC 423
QY 690 AGTCACCTGTTACTGAGTCTGAATACCGGTTGTTTCAAGTTGATCCGGCTTTTGAGGCAA 749
Db 424 AGTCTCTGTACAGGGCTGGAAGTTCCTGTTGTCAGGTTGGATCCAGCCTTCAGGGGT 483
QY 750 TGGATGTTGGAATTTGGAAGCTTGTGTAAGTGGAAACCTGTTTGCATAATGTGAAGGGCCATG 809
Db 484 TGGATGTTGAGGTTGGCAGCTGTGTAAGTGGAAACCTGTTTGCATAATGTGTCAGGGCATGG 543
QY 810 ACATATGGAATCTCTTAGGACCATGATGTCCTGTAATCTTCTCATGCAAAAGAT 869
Db 544 ACATCTGGAATTTCTCTCGTCTATGAAATGGAACCCAGTGGCAATGTGTCAGGGCATGG 603
QY 870 ATGTTTCCATTTGGGTGTGAGCCCTGCACTAGGCTGTTTACCTGGGCAACATGAAAGGG 929
Db 604 ATATTTCAATTTGGGTGTGAGCCATGCACTAGGCTGTACTCTGGCAGCATGAGAGAG 663
QY 930 AAGGAGGTTGTTGTTGGGAGGATGCCAAAGCTAAGGAATGTGTTCTTACAAAGGAATG 989
Db 664 AAGGAAGATGTTGTTGGGAGGATGCCAAAGGCTAAGGAGTGTGGCTTTCATAAAGGAACC 723
QY 990 TAAAGCAGCAAGAGAGGAGTGTATGTAAGTGGCTATCCCATCCCATCCCAATG 1049
Db 724 TCA---AGCAGAGATGGAACAAATAATGGCAATGGCAATG----- 766
QY 1050 GTGATGTACACATGTGCTGACATTTTCAACAGCCGGAATGTAGTTAACTTGAGCAGGA 1109
Db 767 -TGAACCGGCACGTTTCTGATCTCTTCGACACCCAGNNNTGTTACCTTACCCAGGA 825
QY 1110 CTGG 1113
Db 826 CTGG 829

RESULT 9
BI316488
LOCUS
DEFINITION
saf03h04.v1 Gm-cl065 Glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-cl065-4568 5', similar to TR:Q39619 Q39619
PAPS-REDUCTASE-LIKE PROTEIN. ; mRNA sequence.
ACCESSION
BI316488
VERSION
BI316488.1 GI:14990815
KEYWORDS
EST.
SOURCE
Glycine max (soybean)
ORGANISM
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 507)
REFERENCE
1 Shoemaker,R., Keim,P., Vodkin,L., Erpelnding,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
JOURNAL
Contact: Shoemaker R/Public Soybean EST Project
COMMENT
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
```

Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)  
Seq primer: -40RP from Gibco  
High quality sequence stop: 410.  
Location/Qualifiers  
1. 507  
/organism="Glycine max"  
/mol\_type="mRNA"  
/cultivar="Williams"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl065-4568"  
/tissue\_type="germinating shoots"  
/lab\_host="DH10B"  
/clone\_lib="Gm-cl065"  
/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI. The cDNA library was constructed from mRNA isolated from germinating shoots of the cultivar Williams. The seeds were allowed to germinate for 24 hours prior to being cold stressed for 2 days at 4C. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

## ORIGIN

Query Match 30.2%; Score 491.6; DB 3; Length 507;  
Best Local Similarity 98.2%; Pred. No. 9e-125;  
Matches 497; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 21 ACAGCTAGTTAATGGCCCTCGCTTTCACCTTCTTCAATTTCCGACCACTTCCACCTCC 80  
Db 2 ACAGCTAGTTAATGGCCCTCGCTTTCACCTTCTTCAATTTCCGACCACTTCCACCTCC 61

QY 81 CATCATCGAACCACAACTTCCGCAATTTGGTCAATTTAGGATTTCCGAGAGGCCATTG 140  
Db 62 CATCATCGAACCACAACTTCCGCAATTTGGTCAATTTAGGATTTCCGAGAGGCCATTG 121

QY 141 GAGCGCCGTTTAAATTTCAATTTATCTCAAAGACGAGCTTGGTAAAGCCGTTTAAACGCG 200  
Db 122 GATTCTTCTTTTAAATTTCAATTTATCTCAAAGACGAGCTTGGTAAAGCCGTTTAAACGCG 181

QY 201 AACCTCCACGCAAGGATTCATTTGTTCTCTCGAGCAACCACTCGTTCGTTCTGCTT 260  
Db 182 AACCTCCACGCAAGGATTCATTTGTTCTCTCGAGCAACCACTCGTTCGTTCTGCTT 241

QY 261 CTGAGACGAAGAGGAAGATTTTCAACAGATAGCCAGTGTATCTCGACAATGCTTCACCTC 320  
Db 242 CTGAGACGAAGAGAGAGATTTTGAACAGATAGCCAGTGTATCTCGACAATGCTTCACCTC 301

QY 321 TTGAAATCATGGATAGAGCCCTCGACAATTTCCGCAAAATTCGCAACGACATAGCTATTGCTTCAGTG 380  
Db 302 TTGAAATCATGGATAGAGCCCTCGACAATTTCCGCAAAATTCGCAACGACATAGCTATTGCTTCAGTG 361

QY 381 GTGCTGAAGATGTTGCTTTCAGTATCGGAATTTGACGGGTGACCCCTTTTAGGTTT 440  
Db 362 GTGCTGAAGATGTTGCTTTCAGTATCGGAATTTGACGGGTGACCCCTTTTAGGTTT 421

QY 441 TCAGTTTGGACACTGGGAGACTGAAACCCAGAAAATTTTCAACTTTTTCATGCGGTGAGA 500  
Db 422 TCAGTTTGGACACTGGGAGACTGAAACCCAGAAAATTTTCAACTTTTTCATGCGGTGAGA 481

QY 501 AGCATATTGGAATTCGCATTGAGTAC 526  
|||||

Db	482	AGCATTATGGAATTCGCATTGAGTAC	507	
<p>CK264757 959 bp mRNA linear EST 03-AUG-2004 EST710835 potato abiotic stress cdna library Solanum tuberosum cdna clone POABU61 5' end, mRNA sequence. CK264757 CK264757.1 GI:39821735 EST. Solanum tuberosum (potato) Solanum tuberosum Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Solanum. 1 (bases 1 to 959) Buell, C.R., Hart, A., Ziamann, V., Karamycheva, S.A. and Baker, B. Generation of ESTs from abiotic stressed potato tissue Unpublished (2003) Other ESTs: EST710836 Contact: Robin Buell The Institute for Genomic Research 9712 Medical Center Dr, Rockville, MD 20850, USA Email: potato-array@igr.org Clones can be requested from the University of Arizona Genomics Institute via <a href="http://genome.arizona.edu/orders/">http://genome.arizona.edu/orders/</a>. Seq primer: AAT TAG GTG ACA CTA TAG. Location/Qualifiers 1..959 /organism="Solanum tuberosum" /mol_type="mRNA" /cultivar="Kennebec" /db_xref="taxon:4113" /clone="POABU61" /tissue_type="abiotic stress treated leaf and root tissue" /lab_host="DH10B-Tona" /clone_lib="potato abiotic stress cdna library" /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: Solanum tuberosum var. Kennebec plants were grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots:3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cdna library. RNA sample."</p>				
<p>ORIGIN Query Match 29.9%; Score 486.6; DB 7; Length 959; Best Local Similarity 73.1%; Pred.No. 2.6e-123; Matches 691; Conservative 0; Mismatches 224; Indels 30; Gaps 4; QY 509 GGAAATTCGCATTGAGTACATGTTCCCTGATGCTGTTGAGGTTTCAGGCAATTTGAGGAGT 568 Db 1 GGCATTGCGATTGAATACATGTTCCCTGATTCAGTTGAGTTGAGGCTTTGTTAGACC 60 QY 569 AAGGGGTTATCTCTTTCTACGAGGATGGGCACCAAGAGTGTTCAGGGTGAGAAAGGTTG 628</p>				
Db	482	AGCATTATGGAATTCGCATTGAGTAC	507	
<p>RESULT 10 CK264757 LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM REFERENCE AUTHORS TITLE JOURNAL COMMENT FEATURES source 1..959 /organism="Solanum tuberosum" /mol_type="mRNA" /cultivar="Kennebec" /db_xref="taxon:4113" /clone="POABU61" /tissue_type="abiotic stress treated leaf and root tissue" /lab_host="DH10B-Tona" /clone_lib="potato abiotic stress cdna library" /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: Solanum tuberosum var. Kennebec plants were grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots:3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cdna library. RNA sample."</p>				
Db	482	AGCATTATGGAATTCGCATTGAGTAC	507	
<p>ORIGIN Query Match 29.9%; Score 486.6; DB 7; Length 959; Best Local Similarity 73.1%; Pred.No. 2.6e-123; Matches 691; Conservative 0; Mismatches 224; Indels 30; Gaps 4; QY 509 GGAAATTCGCATTGAGTACATGTTCCCTGATGCTGTTGAGGTTTCAGGCAATTTGAGGAGT 568 Db 1 GGCATTGCGATTGAATACATGTTCCCTGATTCAGTTGAGTTGAGGCTTTGTTAGACC 60 QY 569 AAGGGGTTATCTCTTTCTACGAGGATGGGCACCAAGAGTGTTCAGGGTGAGAAAGGTTG 628</p>				
Db	61	AAAGGGCTTTTCTCTTTCTATGAGGATGGCCACCAAGAGTGTTCCTGTGAAGGAAGTT	120	
QY	629	AGGCTTTTAAGGAGGGCCCTTAAGGGTCTCAGAGCATGGATAACTGTGTGAGAGAAAGAC	688	
Db	121	AGGCTTTTGAAGGAGAGCTCTAAAGGGCTTACGTGCTTGGATCACAGGCCAACGTAAAGAT	180	
QY	689	CAGTCACTCTGTTACTAGTCTGAAATACCGGTTGTTTTCAGGTTGATCCGGCTTTTGAAGGA	748	
Db	181	CAGTCCCTCTGGAATCGATCAGAAATCCCATTTGTTTTCAGGTGGACCTCTCTTTTGAAGGA	240	
QY	749	ATGGATGGTGGAAATTTGGAAGCTTGGTGAAGTGAACCCCTCTTGCATATGTAAGGGCCAT	808	
Db	241	TTGGATGGCGGTGCTGGTAGCTTGGTGAATGGAACCCCTGGCTAATGTGGACGGAATA	300	
QY	809	GACATATGGAATCTTCCTTAGGACCATGAATGTGCTGTGAATTCCTTTCATGCAATGCAAAAGGA	868	
Db	301	GATATTTGGAACTTCTGCTGCGTCCATGAATGTGCTGTGAATCATTTGCATTCACAAGGA	360	
QY	869	TATGTTTTCATTTGGGTTGAGCCCTGCATAGGCTGTTTTCAGTGGGCAACATGAAAGG	928	
Db	361	TATGTATCCATTGGATGCGAACTTGCACAAGGCCAGTCTTACGAGGGCAACACGAGAGA	420	
QY	929	GAAGGAGGTGCTGGTGGGAGGATGCCAAAGCTAAGGAATGTGGTCTTTCACAAAGGAAT	988	
Db	421	GAAGGAAGATGGTGGTGGGAAGATGCCAAGGCCAAGGAGTGTGGCTTTGCAAGGGCAAC	480	
QY	989	GTAAGCAGCAGAGAGGAGGAGTGTAAATGGAAATGGGCTATCCCAATCCCATGCAAT	1048	
Db	481	ATCAGGATGAAAGTGTAAATGGCGCT-----GTCCAAACAAT	519	
QY	1049	GGTATGCTACTACCACTGTGCTGCATTTTCAAAGCCCGAATGTAGTTAACTTTGAGCAGG	1108	
Db	520	GGTACTGCTACCATTTG--CTGATATTTTGTATACCAAGGACATTTGTACTTTGAGTAAG	576	
QY	1109	ACTGGAATTCAGAAATTTGGCAAAATTTGGAGACCGAAGAACCACTGCTTTGTGCTTT	1168	
Db	577	CCTGGAGTAGAAGACCTTACTAAAAATTTGGAAGACCGAAGAGAGCTTGGCTCGTTGTTCTT	636	
QY	1169	TATGCACCATGGTCCCTACTCCAGGCTATGAGGAATCTTTATGTTGACTTAGCAGAC	1228	
Db	637	TATGCACCTTGGTGCCAAATTTTCCAGGCAATGGAAGGATCTTATGTTGATTTGGCTGAG	696	
QY	1229	AAGTTAGCAGGGTCAAACAGGGATGAAAGTTGGAAAAATTTAGAGCAGATGGAGAACAGAAA	1288	
Db	697	AGTTGGCTGTTT---CTGGTGTGAAAGTGGGGAATTTACAGGCAGATGGTGACCAAGAAA	753	
QY	1289	GAATTTGCAAGAGTGAATGCAATTTGGGAAGCTTCCCTACGATATATTTTTCACCAAG	1348	
Db	754	GCATTTGCAACAGGAATTTGCAGCTTGGGAGCTTCCCTACAATACTCTTCTTCCCAAG	813	
QY	1349	CATTGCTCTCGGCCAACAAATAAAGTATCCCTCAGAAAAGAGAGATGTTGATTCCTTCATG	1408	
Db	814	CATCTTCAAG---GCCATTAAGTACCTTTCAGAGAAGAGGATGTAGACTCTCTCTGCTG	870	
QY	1409	GCATTTGTAATGCTCTTAAAGATGAGGATATCAGGAAATTTTCTTC	1453	
Db	871	GCCTTTGTGATGCTCTCAGATGAAAGCAAAATGGGTGTTTGATC	915	
<p>RESULT 11 AW508411 LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM AW508411 629 bp mRNA linear EST 24-JUL-2004 si40e11.v1 Gm-r1030 Glycine max cdna clone GENOME SYSTEMS CLONE ID: Gm-r1030-1413 5' similar to TR:Q39619 Q39619 PAPs-REDUCTASE-LIKE PROTEIN. ; mRNA sequence. AW508411 AW508411.1 GI:7146489 EST. Glycine max (soybean) Glycine max Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;</p>				

```

REFERENCE
AUTHORS
  Glycine.
  1 (bases 1 to 629)
  Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corvett, V.,
  Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
  Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
  Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
  Schurr, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
  McCann, R., Waterston, R. and Wilson, R.
  Public Soybean EST Project
  Unpublished (1999)
  Other ESTs: BE823898 corresponding to Gm-r1070-8298 (3')
  Contact: Shoemaker R/Public Soybean EST Project
  Public Soybean EST Project
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: est@watson.wustl.edu
  This clone is available through: ResGen, Invitrogen Corp. 2130
  South Memorial Parkway Huntsville, AL 35801 For further information
  Call: (800) 533-4363 or contact via email: ccu@resgen.com
  Insert Length: 1537 Std Error: 0.00
  High quality sequence stop: 446.

FEATURES
  Location/Qualifiers
    1..629
    /organism="Glycine max"
    /mol_type="mRNA"
    /db_xref="taxon:3847"
    /clones="GENOME SYSTEMS CLONE ID: Gm-r1030-1413"
    /lab_host="DH10B"
    /clone_lib="Gm-r1030"
    /note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This
    cDNA library was constructed from mRNA isolated from
    immature cotyledons of greenhouse grown plants
    (individual seed fresh weight of 100-300mg). The library
    was prepared using the Life Technologies pSuperScript cDNA
    library construction kit. Complementary DNA was
    synthesized from mRNA using a poly(dT) sequence with a
    NotI restriction site. SalI linkers adapters were ligated
    to the blunt-ended cDNA fragments followed by NotI
    digestion. The cDNA fragments were directionally cloned
    into the NotI-SalI restriction site of the pSPORT1
    vector. The ligated cDNA fragments were transformed into
    E. coli ElectroMax DH10B host cells. This library was
    constructed by Dr. Lila Vodkin and Dr. Anu Khanna. Note
    that Gm-r1030 is a re-rack of Gm-cl007."

ORIGIN
  Query Match 29.6%; Score 481.4; DB 1; Length 629;
  Best Local Similarity 85.4%; Pred. No. 6.5e-122;
  Matches 536; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 276 AAGATTTTGAACAGATGCCAGTATCTCGAAATGCTTCACCTCTTGAATCATGGATA 335
DB 1 AAGATTTTGAACAAATTAGCGAAGACCTTGAAATTCATCTCTTGGAGATTATGGATA 60
QY 336 GAGCCCTCGACAAATTCGGCAACGACATAGCTATTTCCTTCAGTGGTCTGAAAGTTG 395
DB 61 AGGCCCTCGAGAAATTTGGGAACGACATCGCTATTTCCTTTAGTGGTCTGAAAGTTG 120
QY 396 CTTTGATTGAGTATGCGAAATTTGACGGTTCGACCCCTTTAGGTTTCAGTTTGGACACTG 455
DB 121 CTTTGATTGAGTATGCAATTTCAGCGGTTCGACCTTCACAGATGTTTATGCTTGACACTG 180
QY 456 GGAGACTGAACCCGAGAACTTATCAACTTTTGTGATCGGTTTGAGAAAGCAATTATGGAATTC 515
DB 181 GGAGACTGAACCCGAGAACTTATCAAAATTTTGTGACGCTTTGAGAAAGCAATTATGGAATTC 240
QY 516 GCATTGAGTACATGTTCCCTGATGCTGTGTAGGTTTCAGGCATTTGGTGGAGTAAAGGGT 575
DB 241 ATATTGAGTACATGTTCCCTGATGCTGTGTAGGTTTCAGGCATTTAGTAAAGCACTAAGGGC 300
QY 576 TATTCTCTTTCTACGAGGATGGCCACCAAGAGTGTGTGACGGTGTGAGAAAGGTGAGGCCTT 635

```

```

301 TCCTCTCATTTTACGAGGATGGGCATCAAGAGTGGCTGTAGAGTAAGAAAGGTGAGGCCCT 360
636 TAGGAGGGCCCTTAAGGCTCTCAGAGCATGGATACTGCTCAGAGGAAAGACCCAGTCCAC 695
361 TGAGGAGAGGCCCTTAAGGGTCTCAAGAGCATGGATTATTCGACAGAGAAAGACCCAGTCTC 420
696 CTGCTACTAGGTTCTGAAATACCCGGTTGTTTCAGGTTGATCCGGCTTTTGGAGGGAATGGATG 755
421 CTGCTACTAGGTTCTGAAATACCCCTATTGTCAGGTTGATCTGTTTGGAGGACTGGATG 480
756 GTGGAATTGGAAGCTTGGTGAAGTGGAAACCTCTGTTGCAAAATGTGAAGGGCCATGACATAT 815
481 GTGGAATTGGAAGCTTGGTGAAGTGGAAACCCGGTTGCAAAATGTTAAATGGTCTGACATAT 540
816 GGAATCTCTTAGGACCATGAATGTGCTGCTGTAATTCCTTGGCAAAAGGATATGTTT 875
541 GGAATCTCTTAGGACCATGAATGTGCTGTAATTCCTTGGCAAAAGGATATGTTT 600
876 CCATTGGGTTGTGAGCCCTGCACTAGGCC 903
601 CGATTGGCTGTGAGCCATGATCAAGGC 628

RESULT 12
CB340603/c
LOCUS
DEFINITION
  CB340603 704 bp mRNA linear EST 14-MAR-2003
  CA23E1011VR_All Cabernet Sauvignon Leaf - CA23E1 Vitis vinifera
  cDNA clone CA23E1011VR_All 3', mRNA sequence.
ACCESSION
  CB340603
VERSION
  CB340603.1 GI:28961185
KEYWORDS
  EST.
SOURCE
  Vitis vinifera
  Vitis vinifera
  ORGANISM
  Vitis vinifera
  Rukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
  rosids; Vitaceae; Vitis.
  1 (bases 1 to 704)
  Goes da Silva, F., Lim H., Iandolino, A., Baek, J., Leslie, A., Xu, J.,
  Jones, K., Walker, M. A. and Cook, D. R.
  Transcriptional responses of Vitis vinifera to infection by the
  bacterial pathogen Xylella fastidiosa
  Unpublished (2003)
  Contact: Douglas Cook, PhD
  CAES Genome Facility
  UC Davis, Plant Pathology
  One Shields Ave, Davis, CA 95616, USA
  Tel: 530 754 6561
  Fax: 530 754 6617
  Email: drcocook@ucdavis.edu
  Seq primer: GCCAAGCAATGGTCTAG.
  Location/Qualifiers
    1..704
    /organism="Vitis vinifera"
    /mol_type="mRNA"
    /cultivar="Cabernet Sauvignon"
    /db_xref="taxon:29760"
    /clones="CA23E1011VR_All"
    /sex="Hermaphrodite"
    /dev_stage="Mid-season leaf material, collected July 25,
    2001."
    /lab_host="DH5alpha"
    /clone_lib="Cabernet Sauvignon Leaf - CA23E1"
    /note="Organ: Leaf; Vector: pDNR; Site 1: SfiI; Site 2:
    SfiI; CA23E1 is a cDNA library of Cabernet Sauvignon
    leaves. The leaves were collected on July 25, 2001, in
    Napa Valley, California, and represent leaves in
    mid-season development. These leaves were verified to be
    infected with the bacterial pathogen, Xylella fastidiosa,
    based on a diagnostic assay using PCR and Xylella-specific
    primer pairs. The plants were asymptomatic at the time of
    collection, but later developed symptoms. cDNAs were made
    by oligo-dT priming and directionally cloned. 5' and 3'

```





Db 339 GAGCTCTTGAAGAGTTTGGAAATGACATAGGATGCTTTTCAGTGGTCTGAAGATGTTG 398  
Qy 396 CTTTGATGTAGTATGCGAAATATGACGGGTGACACCTTTAGGGTTTTCAGTTTGACACTG 455  
Db 399 CTTTGATCGAGTATGCCATTTAACTGCTCGACCATTTAGGGTGTGTTAGCCTGATAGT 458  
Qy 456 GGAGACTGAACCCAGAACATATCACTTTTTCATGCGGTGAGAACATTTATGGAATTC 515  
Db 459 GAAGGCTAAACCCAGAAACATATAGATTCTTTGATGAAGTTTGAGAAACATTTATGGCATAC 518  
Qy 516 GCATTGATGATCATGTTCCCTGATGCTCTGAGGTTTCAGGCATTTGGTGAGGAGTAAGGGGT 575  
Db 519 GTATTGAATACATGTTTCTGATGCTCTTGAAGTTTCAGGCATTTAGTGAGGACAGGGGC 578  
Qy 576 TATTCTCTTTCTAGAGATGGGCAACAGAGTGTTCAGGGTGAAGAGTGAGGCTT 635  
Db 579 TCTTCTCTTCTATAGAGATGGGCACAGGAGTGTGCGGTGAGGAAATGAGGCCCC 638  
Qy 636 TAAGGAGGGCCCTTAAGGTTCTAGAGCATGGATAACTGTTGTCAGAGAAAGACAGTAC 695  
Db 639 TTCGAGGGGCCCTCAAGGGAATTTGGTGCCTGGATCACTGGCCAAAGGAAAGATCAGTCTC 698  
Qy 696 CTGCTACTAGTCTGAATACCGGTTCTTCAAGTTGATCGGCTTTTGAGGGAATGGATG 755  
Db 699 CAGGTACTAGTCTGAATCCCTGTAGTCCAGGTGGATCCTGTTTGGGGCTGGAAG 758  
Qy 756 GTGGAATTTGGAAGCTTGGTGAAGTGGAAACCTGTTGCAAAATGTGAAGGCCCATGACATAT 815  
Db 759 GTGGGTTGGCAGCTGTTGAAGTGGAAATCAGTGGCTTAATGCGAAGGGAATGACATTT 818  
Qy 816 GGAATCTCTTAGACCATGATGTCCTGCTGTAATCTTGCATGCAAAAGGATATCTT 875  
Db 819 GGAATCTCTTCGTAATGATGGGTGTCCTCAATCTTTCACATCAAAAGGGTACATCT 878  
Qy 876 CCATTGGGTGTGACCTCGACTAGGCTGT 906  
Db 879 CGATTGGGTGTGACCATGCAACAGGCCGT 909

RESULT 14  
CF119175/c  
LOCUS  
DEFINITION MTU10CS.P14.D12 Aspen stem cDNA Library Populus tremuloides cDNA,  
mRNA sequence.  
CF119175  
VERSION CF119175.1 GI:33183854  
KEYWORDS EST.  
SOURCE Populus tremuloides (quaking aspen)  
ORGANISM Populus tremuloides  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids I; Malpighiales; Salicaceae; Populus.  
REFERENCE 1 (bases 1 to 764)  
AUTHORS Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and Tsai,C.-J.  
Expressed sequence tags from Aspen  
TITLE Unpublished (2003)  
JOURNAL Contact: Tsai C-J  
COMMENT Plant Biotech Research Center  
Michigan Technological University, School of Forest Resources &  
Environmental Science  
1400 Townsend Drive, Houghton, MI 49931-1295, USA  
Tel: 906 487 2914  
Fax: 906 487 2915  
Email: chtsai@mtu.edu.  
Location/Qualifiers  
1..764  
/organism="Populus tremuloides"  
/mol\_type="mRNA"  
/db\_xref="taxon:3693"  
/clone\_lib="Aspen stem cDNA Library"  
/note="Organ: stem"

ORIGIN  
Query Match 27.7%; Score 450.8; DB 6; Length 764;  
Best Local Similarity 76.1%; Pred. No. 2.1e-113; Indels 6; Gaps 2;  
Matches 583; Conservative 0; Mismatches 177;  
Qy 599 CACCAAGAGTGTTCAGGGTGAAGAGGTGAGCCCTTTTAAAGGAGGCGCTTAAAGGGTCTC 658  
Db 764 CATCAGAGTGTCTCCGTGTGGAAGGTGAGACCCCTTGAGGCGGCTCTGAAGGGGCTT 705  
Qy 659 AGAGCATGGAATACTGGTTCAGAGGAAGACAGTCACTCTGTACTAGTCTGAAATACCG 718  
Db 704 CGGCGCTCGATCACTGGCCAAAGGAAGGATCAATCTCAGGAGCAGGCTCTGAAATTCCA 645  
Qy 719 GTTGTTCAGGTTGATCCGCTTTTGAAGGAATGGATGTGGAATTTGAAAGCTTGGTCAAG 778  
Db 644 GTTGTTCAGGTGACCCGCTTTTGAAGGAATGGATGTGGAAGCTTGAAGCTTATCAAG 585  
Qy 779 TGAACCCCTCTTGCATAATGTGAAGGGCCATGACATATGGAATCTCTTAGGACCATGAAT 838  
Db 584 TGGATCCAATGGCAATGTGAGGACCAAGATCTATGGAAGTTCCTTCGAACCATGGAT 525  
Qy 839 GTGCTGTGAATTCCTTTCATGCAAGAGATATGTTTCCATTTGGGTGTGAGCCCTGCACT 898  
Db 524 GTGCTGTGAATCAITTCATTTCAAAGGGATACATCTCTATTTGGTTGCGAGCCTTGACA 465  
Qy 899 AGGCTGTGTTTACTGCGCAACATGAAGGGAAGGAGGTGTGTTGGAGGATGCCAAA 958  
Db 464 AGGCGAGTTCTTCGGGCCCAACATGAGAGAGAAGAGTGGTGTGGAGGATGCCACA 405  
Qy 959 GCTAAGGAATGTGCTTTCACAAAGGAATGTAAAGCAGCAGAAAGAGGAGATGTTAAT 1018  
Db 404 GCTAAGGAATGTGCTTTCATTAAGGAATCTGAAACAGGGTGATGCAAGCCCACTTAAT 345  
Qy 1019 GGAATGGGTATCCCAATCCCATGCAAAATGGTGATGCTACCACTGTGCTGCAATTTTC 1078  
Db 344 GGCAATGGAACCGGGCTGCCCATGCAAAACGG---GGCTGCCACTGTGTCTGATATTTTC 288  
Qy 1079 AACAGCCCGAATGTAGTTAACTTGAGCAGGACTGGAATTTGAGAAATTTGGCAAAATTTGAG 1138  
Db 287 AACTCCGAGAATTTGGTCAATTTGAGCAGGCCCGGAATCGAAGACTTATTAATATTAGAG 228  
Qy 1139 GACCGAAGGAACCATGGCTTGTGTTGCTTTATGSCACCATGTTGCCCTACTGCCAGCT 1198  
Db 227 AACCGAAGGACCATGGCTTGTGCTCTACGCCCATGTTGCCAATTTTGGCAAGGT 168  
Qy 1199 ATGAGGAATCTTATGTTGACTTAGCAGACAAAGTTAGCAGGTCAACAGGATGAAGTT 1258  
Db 167 ATGGAAGCATCTATGTTGAATTTGGCCGATAAAGTTAGCAGGG---AGTGGAAATAAGGTG 111  
Qy 1259 GGAATAATTAGCAGATGGAGACAGAAAGAAATTTGCAAGAGTGAACCTCAATTTGGGA 1318  
Db 110 GGAATAATTAGCAGATGGAGATGGGATCAGAGGAATTTCAAAACAAAGAAATTAAGCTCGGA 51  
Qy 1319 AGCTTTCCTACGATATTATTTTCCCAAGCATTTGCTCTCGGCCAA 1364  
Db 50 AGTTTCCACACATCTCTTCTTCCCTTAACACATCAITCTCGCCCAA 5  
RESULT 15  
CF119189/c  
LOCUS  
DEFINITION MTU10CS.P14.F03 Aspen stem cDNA Library Populus tremuloides cDNA,  
mRNA sequence.  
CF119189  
VERSION CF119189.1 GI:33183882  
KEYWORDS EST.  
SOURCE Populus tremuloides (quaking aspen)  
ORGANISM Populus tremuloides  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids I; Malpighiales; Salicaceae; Populus.  
REFERENCE 1 (bases 1 to 764)

AUTHORS Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and Tsai,C-J.  
TITLE Expressed sequence tags from Aspen  
JOURNAL Unpublished (2003)  
COMMENT Contact: Tsai C-J  
Plant Biotech Research Center  
Michigan Technological University, School of Forest Resources & Environmental Science  
1400 Townsend Drive, Houghton, MI 49931-1295, USA  
Tel: 906 487 2914  
Fax: 906 487 2915  
Email: chtsai@mtu.edu  
FEATURES  
source location/Qualifiers  
1. .764  
/organism="Populus tremuloides"  
/mol\_type="mRNA"  
/db\_xref="taxon:3693"  
/clone\_lib="Aspen stem cDNA Library"  
/note="Organ: stem"

## ORIGIN

Query Match 27.7%; Score 450.8; DB 6; Length 764;  
Best Local Similarity 76.1%; Pred.No. 2.1e-113;  
Matches 583; Conservative 0; Mismatches 177; Indels 6; Gaps 2;  
QY 599 CACCAAGAGTGTTCAGAGGTGAGAAAGGTGAGGCGCTTTAAGGAGGGCCCTTAAGGGGTCTC 658  
DB 764 CATCAGGAGTGCTGCCGTGTTAGGAGGTGAGACCCCTTGAGGCGGGCTCTGAAGGGGCTT 705  
QY 659 AGAGCATGGATACTGTGTAGAGAGAAAGACAGTACCTGTGTACTAGGTCTGAAATACCG 718  
DB 704 CGGCGCTGGATCACTGGCCAAAGGAGGATCAATCTCCAGGACGAGGTCTGAAATTCGA 645  
QY 719 GTTGTTCAGGTTCATCCGGCTTTTGGAGGAATGGATGGTGAATTCGAAGCTTGGTGAAG 778  
DB 644 GTTGTTCAGGTGACCCGGTTTGGAGGATTTGGATGGTGGGCTGGAACCTGATCAAG 585  
QY 779 TGAACCCCTGTTCGAATGTAAGGGCCATGACATATGGAACCTTCCTTAGGACCATGAAT 838  
DB 584 TGAATCCCAATGCAAAATGTTAGGGACAAAGATGATGGAAGTTCCCTTCGAACCATGGAT 525  
QY 839 GTGCTGTGAATTCCTTGCATGCAAAAGGATATGTTTCCATTTGGGTGTGAGCCCTGCAT 898  
DB 524 GTGCTGTGAATTCATTCGATTCAAAGGGATACATCTCTATTGTTGCCGAGCCTTGCA 465  
QY 899 AGGCTGTTTTTACCTGGCAACATGAAAGGAGGAGGTGTGTGGAGGATGCCAA 958  
DB 464 AGGCCAGTTCTCCGGCCCAACATGAGAGAGAGGAGGTGGTGGGAGGATGCCCA 405  
QY 959 GCTAAGGAATGTGCTCTTCAAAAGGAAATGTAAAGCAGCAGAAAGAGGAGGATGTTAAT 1018  
DB 404 GCTAAGGAATGTGCTCTTCAATGAAGAAATCTGAACAGGGTGATGCAGGCCCACTTAAT 345  
QY 1019 GGAATGGGCTATCCCAATCCCAATGCAAAATGGTGATGTACCACTGTGCTGACATTTTC 1078  
DB 344 GGCAATGGAAACGGGGCTGCCATTCGAAACGG---GGCTGCCACTGTTGCTGATATTTTC 288  
QY 1079 AACAGCCGCAATGTAGTTACTTTCAGCAGGACTGGAATTTGGAATTTGGCAAAATTTGGAG 1138  
DB 287 AACTCGAGAAATTTGGTCAATTTGAGCAGCCCGGAATCGAGAACTTTATTAATAATTAGAG 228  
QY 1139 GACCGAAAGGAACCATGGCTTGTGTGCTTTATGCAACCATGGTGGCCCTACTGCCAGGCT 1198  
DB 227 AACCGAAAGGAACCATGGCTTGTGTGCTCTACGCCCATGGTGGCCAAATTTTCCCAAGGT 168  
QY 1199 ATGGAGGAATCTTATGTTGATCTTAGCAGACAACTTAGCAGGGTCAACAGGATGAAGGTT 1258  
DB 167 ATGGAAGCATCTCTATGTTGAATTTGGCCGATAGTTAGCAGGG---AGTGAATAAAGGTG 111  
QY 1259 GGAATAATTTAGCAGATGAGAACAGAAAGATTTGCCAAGAGTGAACCTGCAATTTGGGA 1318  
DB 110 GGAATAATTTAGGCGATGCGGATCAGAAGGAATTTCTCAAAACAAAGAAATTACAGCTCGGA 51

QY 1319 AGCTTCCCTACGATATATTTTCCCAAGCATTTGTTCTCGGCAA 1364  
DB 50 AGTTTCCCAACAATCCTTTTCTTCCCTTAACACTCATCTCGCCAA 5

Search completed: February 20, 2006, 22:40:40  
Job time : 5984 secs

**THIS PAGE LEFT BLANK**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 20:55:42 ; Search time 314 Seconds  
(without alignments)  
9221.809 Million cell updates/sec

Title: US-10-731-525-7  
Perfect score: 1629  
Sequence: 1 gcacggaggagaccata.....taaaaaaaaaaaaaaaaaaa 1629

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 88878028 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/ptodata/1/ina/1\_COMB.seq:\*
- 2: /cgn2\_6/ptodata/1/ina/5\_COMB.seq:\*
- 3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/ptodata/1/ina/H\_COMB.seq:\*
- 6: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq:\*
- 7: /cgn2\_6/ptodata/1/ina/PP\_COMB.seq:\*
- 8: /cgn2\_6/ptodata/1/ina/RE\_COMB.seq:\*
- 9: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1629	100.0	1629	3	US-09-720-318A-7
2	904	55.5	1795	3	US-09-720-318A-5
3	538.2	33.0	1827	3	US-09-720-318A-9
4	435.2	26.7	1210	3	US-09-720-318A-3
5	364.2	22.4	1215	3	US-09-720-318A-1
6	199.6	12.3	783	3	US-09-328-352-1063
7	149.4	9.2	810	3	US-09-252-991A-13944
8	84.2	5.2	660	3	US-09-252-991A-13837
9	81	5.0	723	3	US-09-902-540-7978
10	81	5.0	6063	3	US-09-902-540-807
11	73.4	4.5	269223	3	US-09-596-002-41
12	72.6	4.5	756	3	US-09-540-236-1037
13	59	3.6	7218	2	US-08-232-463-14
14	42	2.6	1141	3	US-09-806-708B-22
15	39.2	2.4	578	3	US-09-385-982-465
16	38.6	2.4	1830121	3	US-09-557-884-1
17	38.6	2.4	1830121	3	US-09-643-990A-1
18	38.6	2.4	1830121	3	US-10-158-865-1
19	38	2.3	474	3	US-09-621-976-18033
20	36.6	2.2	1101	3	US-08-984-919A-13
21	36.6	2.2	1101	3	US-08-984-919A-14
22	36.6	2.2	1107	3	US-08-781-420-13
23	36.6	2.2	1107	3	US-08-781-420-14
24	36.6	2.2	1107	3	US-08-874-102-13

ALIGNMENTS

RESULT 1

US-09-720-318A-7  
; Sequence 7, Application US/09720318A  
; Patent No. 6730827

; GENERAL INFORMATION:

; APPLICANT: Falco, Saverio Carl

; APPLICANT: Allen, Stephen M.

; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins

; FILE REFERENCE: BB-1167-C

; CURRENT APPLICATION NUMBER: US/09/720,318A

; CURRENT FILING DATE: 2000-12-21

; PRIOR APPLICATION NUMBER: 60/092,833

; PRIOR FILING DATE: 1998-07-14

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 7

; TYPE: DNA

; LENGTH: 1629

; ORGANISM: Glycine max

; US-09-720-318A-7

Query Match	100.0%;	Score 1629;	DB 3;	Length 1629;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1629;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	GCACGAGGAGAACCCATTAACAGCTAGTTAATGGCCCTCGCTTCACTTCTTCAATTTC	60	
Db	1	GCACGAGGAGAGAACCCATTAACAGCTAGTTAATGGCCCTCGCTTCACTTCTTCAATTTC	60	
Qy	61	CGCACCAACTTCCACCTTCCATCATCGGACCCAACTTCCGCAATTTGGTCAATTAG	120	
Db	61	CGCACCAACTTCCACCTTCCATCATCGGACCCAACTTCCGCAATTTGGTCAATTAG	120	
Qy	121	GATTCGAGAGGCCCATTTGGAGGCCGCTTAATTTCAATTATCTCAAGACGGAGCTT	180	
Db	121	GATTCGAGAGGCCCATTTGGAGGCCGCTTAATTTCAATTATCTCAAGACGGAGCTT	180	
Qy	181	GGTAAAGCCGTTTAAAGCCGCACTCCACGAGGATTCATTTGTTCTTCCGAGCAAC	240	
Db	181	GGTAAAGCCGTTTAAAGCCGCACTCCACGAGGATTCATTTGTTCTTCCGAGCAAC	240	
Qy	241	AACCATCGTTGCTTCTGCTTCTGAGAGCAAGAGGAAGATTTTGAACAGATAGCAGTGA	300	
Db	241	AACCATCGTTGCTTCTGCTTCTGAGAGCAAGAGGAAGATTTTGAACAGATAGCAGTGA	300	
Qy	301	TCTCGACAAATGCTTCACTCTTGAATCATGTAGAGCCCTCGACAAATTCGCAACGA	360	
Db	301	TCTCGACAAATGCTTCACTCTTGAATCATGTAGAGCCCTCGACAAATTCGCAACGA	360	



Db 552 AGCAATTAAGAAATCATATTAGTACATGTTCCCTGATCGGTTGAGTTTCAGGCATTAG 611  
Qy 561 TGAGGAGTAAGGGGTTATTTCTTCTTACAGAGATGGCCACCAAGAGTGTTCAGGGTGA 620  
Db 612 TAAGAACTAAGGGGCTCTTCTCATTTTACAGAGATGGGCATCAAGAGTGTCTAGAGTAA 671  
Qy 621 GAAAGGTGAGGCTTTTAAAGAGGGCCCTTAAAGGTTCTCAGAGCATGGAATAACTGGTCAGA 680  
Db 672 GAAAGGTGAGGCTTTTAAAGAGGGCCCTTAAAGGTTCTCAGAGCATGGAATAACTGGTCAGA 731  
Qy 681 GGAAGACCACTGACCTGTTAGTGTCTGAATACCGGTTGTTCAGGTTGATCCGGCTT 740  
Db 732 GAAAGACCACTGCTCTGGTACTAGGTTGAATGCTGAATGCTGAATGCTGAATGCTGGT 791  
Qy 741 TTGAGGAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 800  
Db 792 TTGAGGAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 851  
Qy 801 AGGCGCATGACATATGAACTTCTTAAAGACCATGAATGTCCTGTGAATTCCTTGGCATG 860  
Db 852 ATGCTCTAGACATATGAACTTCTTAAAGACCATGAATGTCCTGTGAATTCCTTGGCATG 911  
Qy 861 CAAAGGATATGTTTCCATTTGGGTTGAGCCCTGCACTAGGCTGTGTTTACCTGGGCAAC 920  
Db 912 CCAAGGATATGTTTCCATTTGGGTTGAGCCCTGCACTAGGCTGTGTTTACCTGGGCAAC 971  
Qy 921 ATGAAAGGAGGAGGTGTTGGTGGAGGATGCAAAAGCTAAGGAATGTGGTCTTCACA 980  
Db 972 ATGAAAGGAGGAGGTGTTGGTGGAGGATGCAAAAGCTAAGGAATGTGGTCTTCACA 1031  
Qy 981 AAGGAAATGTAAGCAGCAAGAGGAGGATGTTAATGAAATGGGCTATCCCAATGCC 1040  
Db 1032 AAGGAAATGTAAGCAGCAAGAGGATGTTAATGAAATGGGCTATCCCAATGCC 1091  
Qy 1041 ATGCAATGTGTATGCTACCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1100  
Db 1092 A-----TGGCTCTGCCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1142  
Qy 1101 TGAGCAGGACTGGAATTTGGAATTTGGCAAAATTTGGAGGACCGAAAGGAACCAATGGCTTG 1160  
Db 1143 TGAGCAGGCTCGGATTTGAGAAATTTGGCAAAATTTGGAGGACCGAAAGGAACCAATGGCTTG 1202  
Qy 1161 TTGTGCTTTATGACCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1220  
Db 1203 TTGTGCTTTATGACCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1262  
Qy 1221 TAGCAGCAAGATTAGCAGGCTCAACAGGATGAAAGTTGAAATTTAGAGCAGATGGAG 1280  
Db 1263 TGGCAGAGAAATTAGCAAGGT---CAGGAGTGAAGGTTGCAAAATTTAGAGCAGATGGAG 1319  
Qy 1281 AACAGAAAGAAATTTGCAAGAGTGAATGCAATTTGGGAAGCTTCCCTACCATATATTTT 1340  
Db 1320 AGCAGAAAGAAATTTGCAAGAGTGAATGCAATTTGGGAAGCTTCCCTACCATATATTTT 1379  
Qy 1341 TCCCAAGCACTGCTCTCGGCGCAACATAAAGTATCCCTCAGAAAAGAGAGATGTTGATT 1400  
Db 1380 TCCCAAGCACTGCTCT---CAACCAATTAAGTACCCTTCCCAAGAGAGATGTTGATT 1436  
Qy 1401 CTTGATGGCAATTTGAAATGCTTAAAGATGAGAT 1436  
Db 1437 CATTGACGGCAATTCGTAATGCTTACGGTATGGT 1472

## RESULT 3

US-09-720-318A-9  
; Sequence 9, Application US/09720318A  
; Patent No. 6730827  
; GENERAL INFORMATION:  
; APPLICANT: Falco, Saverio Carl  
; APPLICANT: Allen, Stephen M.  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167-C  
; CURRENT APPLICATION NUMBER: US/09/720,318A

; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 1998-07-14  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 9  
; LENGTH: 1827  
; TYPE: DNA  
; ORGANISM: Triticum aestivum  
US-09-720-318A-9

Query Match 33.0%; Score 538.2; DB 3; Length 1827;  
Best Local Similarity 68.1%; Pred. No. 5,7e-152;  
Matches 801; Conservative 0; Mismatches 358; Indels 18; Gaps 3;

Qy 257 GCTTCTGAGACGAAAGGAAGATTTTGAACAGATAGCCAGTGTCTTCGCAATGCTTCA 316  
Db 323 GCATCTGCGTGGCGCGCTGGACTAGAGGCCCTGGCGAGGAGCTGTGGGCGCGCTCG 382  
Qy 317 CCTCTTGAATCATATGATAGAGCCCTCGACAAATTCGGCAACGACATAGCTATTGCTTC 376  
Db 383 CGCTGGAGATCATGATCGTGGCTCGACATGTTCGGCTCCGAAATCGCCATCGCCTTC 442  
Qy 377 AGTGTGCTGAAGATGTTGCTTGTATGATGATGCGAAATTTGACGGGTGCAACCTTTAGG 436  
Db 443 AGTGTGCTGGAGACGCTGGCCCTCATCGAATACGCAAACTGACTGGACGCCCTTCAGG 502  
Qy 437 GTTTTCAGTTTGGACACTGGGACACTGMAACCCAGAACTTATCAACTTTTGTATGCGGTT 496  
Db 503 GTTTTCAGCTTGGACACTGGGACACTGMAACCCAGAACTTATCAACTTTTGTATGCGGTT 562  
Qy 497 GAGAAGCACTATGGAATTCGCATTTGAGTATCATCTTCCCTGATGCTGTGAGGTTTCAGGCA 556  
Db 563 GAGAAGCACTATGGAATTCGCATTTGAGTATCATCTTCCCTGATGCTGTGAGGTTTCAGGCA 622  
Qy 557 TTGCTGAGGAGTAAGGGGTTATTTCTTCTTACAGAGTGGGCAACCAAGAGTGTTCAGG 616  
Db 623 CTTGTGAGGAGTAAGGGGTTATTTCTTCTTACAGAGTGGGCAACCAAGAGTGTTCAGG 682  
Qy 617 GTGAGAAAGTGAAGGCTTTAAGGAGGCGCTTAAAGGCTCTAGAGCATGATGATTAACGTT 676  
Db 683 GTGAGAAAGTGAAGGCTTTAAGGAGGCGCTTAAAGGCTCTAGAGCATGATGATTAACGTT 742  
Qy 677 CAGAGAAAGACCCAGTCACTGCTAGTGTGTAATACCGGTTGTTTCAGGTTTCAGGCG 736  
Db 743 CAGAGAAAGACCCAGTCACTGCTAGTGTGTAATACCGGTTGTTTCAGGTTTCAGGCG 802  
Qy 737 GCTTTTGAAGGAAATGGAATTTGAAAGCTTGGTGAAGTGGAAACCTCTGTTCGAAAT 796  
Db 803 TCTTTTGAAGGAAATGGAATTTGAAAGCTTGGTGAAGTGGAAACCTCTGTTCGAAAT 862  
Qy 797 GTGAGGCGCATGACATATGGAATTTCTTGTAGGACCATGATGCTGCTGTGAATTCCTTG 856  
Db 863 GTGAGGCGCATGACATATGGAATTTCTTGTAGGACCATGATGCTGCTGTGAATTCCTTG 922  
Qy 857 CATGCAAAAGGATATGCTTTCCATTGGTGTGAGCCCTGCACTAGGCTGTTCACCTGGG 916  
Db 923 CATGCTCAAGGCTATGCTTCCATTGGTGTGAGCCCTGCACTAGGCTGTTCACCTGGG 982  
Qy 917 CAAATGAAAGGAGGAGGAGTGGTGGGAGGATGCGCAAAAGCTAAGGAATGTGTCCTT 976  
Db 983 CAGCAGAGAGGAGGAGGAGTGGTGGGAGGAGCGCCACCAAGGAGTGGGCTG 1042  
Qy 977 CACAAAGGAAATGTAAGCAGCAGAAAGAGGAGGATGTTAATGGAATGGGCTATCCAA 1036  
Db 1043 CACAAAGGAAATGTAAGCAGCAGAAAGAGGAGGATGTTAATGGAATGGGCTATCCAA 1090  
Qy 1037 TCCCATGCAAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 1096  
Db 1091 AAGGCAACGGCTCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1150  
Qy 1097 AACTTCAGCAGGAGTGGAAATTTGGAATTTGGCAAAATTTGGAGGAGCGGAAAGCAATGG 1156



Db 1151 AACCTACCCGTCGGGATCGAGAACTCTCTCGGCTCGAGAACCGCGCGAGCCGTGG 1210  
Qy 1157 CTTGTTGTGCTTTANGACCAATGTTGCCCCCTACTGCGAGGCTATGAGGAATCTTATGTT 1216  
Db 1211 CTACCCGTCCTCTACGCTCCCTGCTGTCCTACTGCGAGGCAATGAGGCGCTCTACGTT 1270  
Qy 1217 GACTTAGCAGACAAGTTAGCAGGCTCAACAGGATGAAGTTGGAATTTAGAGCAGAT 1276  
Db 1271 GAGTCGCGGAGAGCTGAGCGGT---CAGGCATCAAGGTGGCCAAAGTTCCGCGCGGAC 1327  
Qy 1277 GGAGAACAGAAAGAAATTTGCAAGAGAGTGAACCTGCAATTTGGGAAGCTTCCCTACGATATTA 1336  
Db 1328 GCGAGCAGAGCCATTCGCGCAGCGGAGCTGCAACTACAGAGCTTCCCGACGATCCTC 1387  
Qy 1337 TTTTCCCAAGCAATTCGCTCTCGGCAACAATAAGATATCCCTCAGAAAGAGAGATGTT 1396  
Db 1388 CTGTTCCCGCGCCGACCGTGAAGC---CCATCAAGTACCCGTCGAGAGAGGAGCGTC 1444  
Qy 1397 GATTCCCTTGATGGCAATTTGTAATGCCCTTAAGATGAG 1433  
Db 1445 CAGTCCCTCTCGGCTTCGTGAACAGCCTCAGATGAG 1481

## RESULT 4

US-09-720-318A-3  
; Sequence 3, Application US/09720318A  
; Patent No. 6730827

; GENERAL INFORMATION:  
; APPLICANT: Falco, Saverio Carl  
; APPLICANT: Allen, Stephen M.  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167-C  
; CURRENT APPLICATION NUMBER: US/09/720,318A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 1998-07-14  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 3  
; LENGTH: 1210  
; TYPE: DNA  
; ORGANISM: Impatiens balsamia  
US-09-720-318A-3

Query Match 26.7%; Score 435.2; DB 3; Length 1210;  
Best Local Similarity 69.7%; Pred. No. 6.5e-121;  
Matches 641; Conservative 0; Mismatches 258; Indels 21; Gaps 3;

Qy 516 GCATTGAGTACATGTTCCCTGATGCTGTTGAGGTTTCAGGCATTTGTTGAGGAGTAAGGGGT 575  
Db 1 GCACGAGGTACATGTTCCCTGATGCAATTTAGTACAAGGATTAGTAAGAACCAAGGAC 60  
Qy 576 TATTCTCTTTCTACGAGGATGGCAACCAAGAGTGTTCAGGGGTGAGAAAGGTGAGGCCCTT 635  
Db 61 TGTCTCTTTCTACGAAGACGACATCAAGAGTGTGCGCGTCAAGAAAGTGAAGCCAC 120  
Qy 636 TAAGAGAGGCCCTTAAGGGTCTCAGAGCATGATTAATCTGTCAGAGGAAAGACCAAGTCAC 695  
Db 121 TGAGCGGTGCTCTCAAGGGTCTCCGCGCTTGATCAAGGATGATGAGGAGGAGGAGGAG 180  
Qy 696 CTGGTACTAGTCTGAATACCGGTTGTTGAGGTTGATCGGCTTTTCAGGGAATGGATG 755  
Db 181 CGGGAAACAGATCGGAGATCCAGTCGTCAGTGGATCCCTCTTTTGAAGGATTTGGTTG 240  
Qy 756 GTGGAATTTGGAAGCTTTGTTGAAGTGGAAACCCCTGTTTGAATGTTGAAGGCCCATGACATAT 815  
Db 241 GTGGAGAGGTTAGCTTGTGTAAGTGGATCCGCTGCTTAAGTAGATGCTCTGTATGAT 300  
Qy 816 GGAATCTCTTAGGACCAATGATGCTGCTGGAATTCCTTGATGCAAAAGATATGTTT 875  
Db 301 GGAATTTCTCCGAGCTATGAATGTCCTGTTAATGCACCTTCATAGCCAGGGTTATGTCT 360  
Qy 876 CCATTGGGTGTGAGCCCTGCACTAGGCTGTTTTCCTAGTGGGCAACATGAAGGGAAGGA 935

Db 361 CGATTGGTTCGGAACCGTGCACCCGCGGTGTACTCTGGCAACATGAGAGAGAGCA 420  
Qy 936 GGTGGTGGTGGAGGATGCAAGCTTAAGGAATGTGTCTTCAAAAGGAATTAAGC 995  
Db 421 GGTGGTGGTGGAGGATGCTGCGGCTTAAGGAGTGTGGCTTACATAAAGGAATATAA--- 477  
Qy 996 AGCAGAAAGAGGAGGATGTTAATGGAATGGGCTATCCCAATCCCATGCAAAATGTTGATG 1055  
Db 478 -----AGGATGCCAATGGGAATGGGGTTGCTCAAGCTGAGGGAGGAAGAA 525  
Qy 1056 CTACCACTGTGCTCTGACATTTTCAACAGCCCGAATGTAGTTAACTTGAGCAGGACTGAA 1115  
Db 526 CTGTTACGGATGCTGATATTTTGAATCCAAGATGTGTGACACTGAGTAGAAGCGGA 585  
Qy 1116 TTGAGAATTTGGCAAAATTTGAGAGCAGCAAGGAACCATGGCTTGTGTGCTTTATGAC 1175  
Db 586 TTGAGAATCTGTGCAAACTTTCAGGAGAGGAAGAGCCATGGATCGTGTCTGTATGAC 645  
Qy 1176 CANTGTGCCCTACTGCGCAGGCTATGAGGAATCTTATGTTGACTTAGCAGACAAAGTTAG 1235  
Db 646 CTTGGTCCAGTCTCTGCCAGGTAATGGAATAATCTACTTGGAAATTTGCTGAAAGCTGG 705  
Qy 1236 CAG---GGTCAACAGGATGAAGGTTGGAAAAATTTAGAGCAGATGAGAAACAGAAAGAA 1292  
Db 706 CGGTGACCGGTGGTGTGAAGGTAGGAAATTTCCGGCAGATGTTGCAGAAAAGAGT 765  
Qy 1293 TTGCAAGAGTGAATCGAATTTGGGAAGCTTCCCTACGATATTAATTTTCCCAAGCAAT 1352  
Db 766 TTGCTACCAAGAATTTGAGCTGGGAGCTTTCACAACTACTCTTCTTCCCAACACT 825  
Qy 1353 CGTCTCGGCCAACAAATTAAGTATCCCTCAGAAAGAGAGATGTTGATTCCTTGATGGCAT 1412  
Db 826 CATCT---AAAGCCATCAAGTACCATCTGAGAAAAAGGAGCGTGGAGTCATTGTTGGCTT 882  
Qy 1413 TTGTAATGCTTAAAGATGA 1432  
Db 883 TTGTGAACGCACTCAGATGA 902

## RESULT 5

US-09-720-318A-1  
; Sequence 1, Application US/09720318A  
; Patent No. 6730827

; GENERAL INFORMATION:  
; APPLICANT: Falco, Saverio Carl  
; APPLICANT: Allen, Stephen M.  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167-C  
; CURRENT APPLICATION NUMBER: US/09/720,318A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 1998-07-14  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 1  
; LENGTH: 1215  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (273)  
; OTHER INFORMATION: n = a, c, g or t  
US-09-720-318A-1

Query Match 22.4%; Score 364.2; DB 3; Length 1215;  
Best Local Similarity 65.2%; Pred. No. 2e-99;  
Matches 591; Conservative 0; Mismatches 294; Indels 21; Gaps 3;

Qy 528 TGTTCCTCTGATGCTGTTGAGGATTCAGGCATTTGAGGAGTAAGGGTTATCTTCTTCT 587  
Db 1 TGTTCCTCGAGCGCAGGAGGCTGCGAGGAGTGGTGGCGCAACCAAGGGCTCTTCTCTTCT 60



```
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13944

Query Match      9.2%; Score 149.4; DB 3; Length 810;
Best Local Similarity 52.2%; Pred. No. 1.6e-34;
Matches 361; Conservative 0; Mismatches 321; Indels 9; Gaps 1;

Qy 302 CTGACAAATGCTTACCTCTTGAATCATGATAGAGCCCTCGACAAATTCGGCAAGCAC 361
Db 115 CTTGCCGACAAATCCCCCGAGGACATCTTGAAGCCGCTTCGAGACATTCCTGGCGAG 174
Qy 362 ATAGCTATTGCTTCAGTGTGCTGAAGATGTTGCTTTGATTGATGATGCGAAATTCACG 421
Db 175 CTGTGATCTCTTACGCGGCGCGAGGAGCTGCTCTAGTAGACATGGCTGGAAGCTC 234
Qy 422 GGTGACCCCTTTAGGGTTCAGTTTGGACACTGGGAGACTGAAACCCAGAAATTTATCAA 481
Db 235 AACCGCAACGTGAAGTGTTCAGCTTCGACACACCGGTGCGCTGCACCCGGAACCTATCGT 294
Qy 482 CTTTTGATCGGTTGAGAGCAATATGGAATTCGCAATTCGATTCAGTACATGTTCCCTGATGCT 541
Db 295 TTTCATGACACGAGTCCGCGAACACTACGGCATCGCCATCGACGTGCTTCCCCCGACCG 354
Qy 542 GTTGAGGTTGAGGCAATTTGGTGAGGAGTAAGGGGTATTTCTTTCTACGAGGATGGGCAC 601
Db 355 CGCTCTGTCGACCCCTGTGTGAAGAAAGGGCCCTGTTACGTTCTACCGGAGCGGCAC 414
Qy 602 CAAGAGTGTTCGAGGTTGAGAAAGGTGAGGCTTTTAAAGGAGGCGCTTAAAGGTTCTCAGA 661
Db 415 GGTGAGTGTGCGGCATCCGCAAGATCGAACCGCTCAAGCGCAAGCTCGCGCGGTGCGC 474
Qy 662 GCATGATTAATGCTGACAGAAAGACAGTCACTGCTACTAGTCTGTAATACCGGTT 721
Db 475 GCCTGGGCGCACCGGCCAACCGCGCACGAGAGCCCGCGCACGCGCAGCGAGTGGCGGTG 534
Qy 722 GTTCAGGTTGATCGGCTTTTGGGGAATGATGTTGGAATTTGGAAGCTTGGTCAAGTGG 781
Db 535 CTGGAATTCGAGGTGCTTTTCCACGCCGA-----AAAGCCGCTGTACAAATTC 585
Qy 782 AACCTGTGTGCAAAATGTGAAGGGCCATGACATATGGAATTCCTTTAGAACCATGAATGTG 841
Db 586 AACCCACTGCTGTCATGATGACGAGGAGGTCTGGGGCTATATCCGATGCTCGAACTG 645
Qy 842 CCTGTGAATTCCTTGCATGCAAAAGGATATGTTTCCATTTGGTGTGACCTGCACTAGG 901
Db 646 CCCTACACAGCTGTCAGAAACGCGCTATATCAGCATCGGCTCGAACCCCTGACCCGT 705
Qy 902 CCTGTTTACCTGGCAACATGAAGGAGGAGGTGTTGTTGGGAGGATGCCAAAGCT 961
Db 706 CCGTGTCTGCCAACACGACGAGCGGAGGCGCGCTGTTGTTGGGAGGAGCCACCCAC 765
Qy 962 AAGGAATGTGCTTTCACAAAGGAATGTAA 992
Db 766 AAGGAGTGGGGTCTGACGCGCGCAACCTGA 796

RESULT 8
US-09-252-991A-13837
; Sequence 13837, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
```

```
; SEQ ID NO 13837
; LENGTH: 660
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13837

Query Match      5.2%; Score 84.2; DB 3; Length 660;
Best Local Similarity 54.8%; Pred. No. 7.4e-15;
Matches 167; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

Qy 302 CTGACAAATGCTTACCTCTTGAATCATGATAGAGCCCTCGACAAATTCGGCAAGCAC 361
Db 356 CTTGCCGACAAATCCCCCGAGGACATCTTGAAGCCGCTTCGAGACATTCCTGGCGAG 415
Qy 362 ATAGCTATTGCTTCAGTGTGCTGAAGATGTTGCTTTGATTGATGATGCGAAATTCACG 421
Db 416 CTGTGATCTCTTTCAGCGCGCGCGGAGGACGTGCTCTAGTAGACATGGCTGGAAGCTC 475
Qy 422 GGTGACCCCTTTAGGGTTCAGTTTGGACACTGGGAGACTGAAACCCAGAAATTTATCAA 481
Db 476 AACCGCAACGTGAAGTGTTCAGCTTCGACACACCGGTGCGCTGCACCCGGAACCTATCGT 535
Qy 482 CTTTTGATCGGTTGAGAAAGCAATATGGAATTCGCAATTCGATTCAGTACATGTTCCCTGATGCT 541
Db 536 TTTCATGACACGAGTCCGCGAACACTACGGCATCGCCATCGACGTGCTTCCCCCGACCG 595
Qy 542 GTTGAGGTTGAGGCAATTTGGTGAGGAGTAAGGGGTATTTCTTTCTACGAGGATGGGCAC 601
Db 596 CGCTCTGTCGAAACCCCTGGTGAAGAAAGGGCTGTTTACGTTCTTACCGGAGCGGCAC 655
Qy 602 CAAGA 606
Db 656 GGTGA 660

RESULT 9
US-09-902-540-7978
; Sequence 7978, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 7978
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-7978

Query Match      5.0%; Score 81; DB 3; Length 723;
Best Local Similarity 49.6%; Pred. No. 7.3e-14;
Matches 346; Conservative 0; Mismatches 315; Indels 36; Gaps 4;

Qy 292 AGCCAGTGTATCTGACAAATGCTTCACTCTTGAATCATGATAGAGCCCTCGACAAATTC 351
Db 42 AGCCGCTGAGTGAAGAGCGCCCGCGGAGACCTCTCTGCTGGACCGAGCGCGCTT 101
Qy 352 CGGCAACGACATAGCTATTGCTTTCAG---TGGTGTGGAAGATGTTGTTGATTGATGA 408
Db 102 CGGTGCGCGCGCGGCCCATGGCTCCAGCTTCGGGTGGAGGACATGTTCTTCATCGACCT 161
Qy 409 TGGCAATTTGACGGTGCAGCCCTTTAGGTTTTCAGTTTTCAGACTGGGAGACTGAAACC 468
Db 162 GCGCGCGCAGCATGCGCCCGACCTTTCAGCGTCTTTCACGCTCGACACCGGACGCTGCC 221
```

469	Qy	AGAACTTATCAACTTTTGTGATCGGTTGAGAA	GCATTATGGAATTCGATTTGAGTACAT	528
222	Db	GGAGACGTACGAACTCATGGAGGTGGTGC	GTGTAAGCGCTACCGCGTACCGTGGAGACGTA	281
529	Qy	GTTCCTTGATGCTGTGTGAGGTTTCAGGCA	ATTCGTCGAGGAGTAAAGGGTTATTTCTCTTTTCTA	588
282	Db	CTTCCCCGAGCGCGCGGTGGAGCGTTGG	AGTTCACGAAACGGCTACTTCTCTCTTCCTCGG	341
589	Qy	CGAG-----GATGGGACCAAGAGTGT	TGCAGGCGTGAGAAAGGTGAGGCGCTTTTAAGGAG	642
342	Db	CCAGAGCTAGAGGCACCGAAGGCGTGTG	CGCCATCCGCAAGGTGGAGCCCGCTGTCCGG	401
643	Qy	GGCCCTTAAGGGTCTCAGAGCATGATAA	CTGGTTCAGAGGAAAGACCAAGTCA	702
402	Db	CGCGCTCGCGGTTCAACAGGCGTGGGT	GACGGGATTTGCGCCGTGAGCAGT---CGGTCA	458
703	Qy	TAGTCTGAAATACCGGTTGTTACAGTT	GATCCGCGCTTTTTCAGGGGAATGGATGGTGGAA	762
459	Db	CCGCACCGACGTTGCCAGCGTTTAGAG	GTGGACCGCGCATG-----	499
763	Qy	TGGAAGCTTTGGTGAAGTGGAAACCTG	TGTGCAAAATGTGAAGGGCCATGACATATG	822
500	Db	-----GGCTGCTCAAGCTCAACCG	CTGCGCCACGTTGGAGCAGCCGCGACATCTG	554
823	Qy	CTTTAGGACCATGAATGTGCTGTGAAT	TCTTCGATGCAGAAAGATATGTTTTCATTGG	882
555	Db	CGTCCGGGCGAAGAGCGTGCCCTACA	ACGCGCTCCATGACCGTGGCTACCCGTC	614
883	Qy	GTGTGAGCCCTGCATCTAGGCGCTG	TTTACCTGGGCAACATGAAGGGGAAGGGAGTGG	942
615	Db	CTGTGCGCCCTGCACGCGCGCGGT	GAAACCCCTACGAGACGAGCGCGCGGCG	674
943	Qy	GTGGGAGGATGCCAAAGCTAAGGAAT	GTGGTCTTCAC	979
675	Db	GTGGGAGTCCGCGAGAAACCGCAG	TGGCGCTCCAC	711

RESULT 10

```

US-09-902-540-807/C
; Sequence 807, Application US/09902540
; Patent No. 683447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: MYXOCOCCUS XANTHUS Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 807
; LENGTH: 6063
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-807

```

Query Match	5.0%;	Score 81;	DB 3;	Length 6063;
Best Local Similarity	49.6%;	Pred. No. 2.5e-13;		
Matches 346;	Conservative 0;	Mismatches 315;	Indels 36;	Gaps 4;

  

Qy	292	AGCCAGTGATCTCGACAATGTTCACTCTTTGAAATCATGGATAGAGCCCTCGACAAATT	351
Db	3803	AGCCGTGAAGTGAAGACGCCCGCGCAGGACCTCTCGCTGGACCGAGCGCGCGCTT	3744
Qy	352	CGGCAACGACATAGCTATTGCTTTCAG---TGGTGTGAAGATGTTGCTTTGATTGAGTA	408
Db	3743	CGGTGCGCGCGGCCCATCGCTCCAGCTTCGGCGTGAGAGACATGTCCTCATCGACCT	3684
Qy	409	TGGGAATTGACGGGTGACCCCTTACGGTTTTCAGTTTGGACACTGGGAGACTGAACCC	468

Db	3683	GGCGCGCCAGCATGCGCCCAAGCTGGCGCTCTTCCACGCTCGAACCGCTGCGCCCTGCCCC	3624
Qy	469	AGAACTTATCAACTTTTTCATCGGTTGAGAAAGCAATATGGAATTGCGAATTGAGTACAT	528
Db	3623	GGAGACGTACGAACCTCATGAGGTGGTGGTAAGCGCTACCGCTGGAGACGTA	3564
Qy	529	GTTCCCTGATGCTGTGAGGTTCAAGGCATTTGGTGAGGAGTAAGGGGTATTTCTCTTTCTA	588
Db	3563	CTTCCCCGAGCGCGCGCTGGAGGCGTTGGAGTCCACGAACGGCTACTTCTCCTTCCG	3504
Qy	589	CGAG-----GATGGGCACCAAGAGTGTCCAGGTTGAGAAAGTGAGGCTTTAAGGAG	642
Db	3503	CCAGGCTTAGAGGCACGCAAGGCGTCTGCGCCATCCGCAAGGTGGAGCCCTGTGCGG	3444
Qy	643	GGCCCTTAAAGGTTCTCAGAGCATGGATAACTGCTCAGAGGAAAGACAGATCAGCTGGTAC	702
Db	3443	CGGCTCGCGGTTCAACAGCGCTGGTGACGGGATTTGCGCGGTGAGCAGT---CCGTAC	3387
Qy	703	TAGGTTGAAATACCGGTGTTTCAGGTTGATCCGGCTTTTGGAGGAATGATGGTGGAA	762
Db	3386	CGCACCGAGCTGGACGTTTAGAGGTGGACAGCGGCATG-----	3346
Qy	763	TGGNAGCTTGGTGAAGTGGAAACCTGTTGCAATATGTAAGGGCCATGACATATGGAAC	822
Db	3345	-----GGCTGCTCAAGCTCAACCGCTGGCCACGTGGAGCAGCGGGACATCTGGGCCTA	3291
Qy	823	CCTTAGACCATGAATGTGCTGTGAATTCCTTGGCATGCAAAAGGATATGTTTCCATGG	882
Db	3290	CGTCCGGCGCAAGAGCGTGCCCTTAACGCGCTTCCATGACCGTGGCTACCCGTCCTATCGG	3231
Qy	883	GTGTGAGCCCTGCATAGGCCTCTTTTACCTGGGCAACATGAAGGGAAGGAGGTGGTG	942
Db	3230	CTGTGCGCCCTGCACGCGCGCGGTGAACCTACGAGGACGAGCGCGGCGCGCTGGTG	3171
Qy	943	GTGGGAGATGCGAAAGCTAAGGAATGTGCTTTCAC	979
Db	3170	GTGGAGTCCGGAGGAACCGCAGTGGCGCTCCAC	3134

RESULT 11

```

US-09-596-002-41
; Sequence 41, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Patterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCES: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,121
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 41
; LENGTH: 269223
; TYPE: DNA
; ORGANISM: Moraxella catarrhalis
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte template ID No. 6632636 41
; PUBLICATION INFORMATION:
US-09-596-002-41

```

Query Match 4.5%; Score 73.4; DB 3; Length 269223;  
Best Local Similarity 58.4%; Pred. No. 4.5e-10;  
Matches 128; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

Qy 771 TGTGTAAGTGGAAACCTGTTTGCAATGTGAAGGCCATGACATATGGAACCTTCCTTAGGA 830  
Db 233373 TTGCCAAATACAAATCCAAATTTTTCATTTGGCAAGACACGATGTATGGGCATATATTTTGA 233432

```
Qy 831 CCATGATGCGCTGCTGAATTCCTTGTCATGCAAGAGATATGTTCCATTGGGTGAGC 890
Db 233433 CCAAAATATACCGTTTAATGAGCTGTATCACCAGGCTATCCTTCCATTGGGTGAGC 233492
Qy 891 CCTGCACTAGGCTGTTTACCTGGGCAACATGAAAGGAGGAGGTGGTGGGAGG 950
Db 233493 CTGTACCATCGCTGTCAGCAAGGCGAGGATATCCGTGCAGGACGCTGGTGGTGGGAGC 233552
Qy 951 ATGCCAAGCTAAGGAATGCTTTCACAAAGGAATG 989
Db 233553 ATAAAGATAAAGAGTGGGTTTACACAAATGACTTG 233591

RESULT 12
US-09-540-236-1037
; Sequence 1037, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540.236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 1037
; LENGTH: 756
; TYPE: DNA
; ORGANISM: M.catarrhalis
US-09-540-236-1037

Query Match 4.5%; Score 72.6; DB 3; Length 756;
Best Local Similarity 58.6%; Pred. No. 2.6e-11;
Matches 126; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

Qy 771 TGGTGAAGTGAACCCCTGTTGCAAAATGTGAAGGCCCATGACATATGGAATTCCTTAGGA 830
Db 542 TTGCAAAATACATCCAAATTTGATGTCGACAGACGATGATAGGCATATATTTGA 601
Qy 831 CCATGAATGTCCTGTGAATTCCTTGCATGCAAAAGGATATGTTTCCATTGGGTGAGC 890
Db 602 CCAAAATATACCGTTTAATGAGCTGTATCACCAGGCTATCCTTCCATTGGCTGTGAGC 661
Qy 891 CCTGCACTAGGCTGTTTACTCGGCAACATGAAAGGAGGAGGTGGTGGGAGG 950
Db 662 CTGTACCATGCTGTCAAGCAAGCGAGGATATCCGTGCAGGACGCTGGTGGTGGGAGC 721
Qy 951 ATGCCAAGCTAAGGAATGCTTTCACAAAGGA 985
Db 722 ATAAAGATAAAGAGTGGCGTTTACACAAATGA 756

RESULT 13
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOMLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ5gt-Fls
US-08-232-463-14

Query Match 3.6%; Score 59; DB 2; Length 7218;
Best Local Similarity 1.3%; Pred. No. 1.3e-06;
Matches 5; Conservative 230; Mismatches 140; Indels 0; Gaps 0;

Qy 924 AAAGGAAGGAGGTGGTGGGAGGATGCCAAGCTAAGGATGTGGTCTTCACAAAG 983
Db 1434 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1375
Qy 984 GAAATGTAAGCAGCAGAAAGAGGAGGATGTTAATGGAATGGGCTATCCCAATCCCATG 1043
Db 1374 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1315
Qy 1044 CAAATGTGTATGCTACCATGTGCTGACATTTTCAACAGCCGCAATGTAGTTAATTGA 1103
Db 1314 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1255
Qy 1104 GCAGGACTGGAATTGAGAAATTTGCAAAATTTGGAGGACCGAAGAACCATGGCTGTGTG 1163
Db 1254 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1195
Qy 1164 TGCTTTATGCACCATGGTCCCTCTACTGCCAGGCTATGGAGGAATCTTATGCTAGTAG 1223
Db 1194 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1135
Qy 1224 CAGACAAAGTTAGCAGGCTCAACAGGATGAAGGTTGAAATTTAGACGACATGGAGAAC 1283
Db 1134 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1075
Qy 1284 AGAAAGAAATTTGCAA 1298
Db 1074 RRRRRRRRATGCGCAA 1060

RESULT 14
US-09-806-708B-22
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
```

```

RESULT 15
US-09-385-982-465/c
; Sequence 465, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS: II
; FILE REFERENCE: CCDNA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; CURRENT FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 465
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

**LEFT BLANK**

**THIS PAGE LEFT BLANK**





Qy 301 TCTCGAATGCTTCACTCTTGAAATCATGGATAGAGCCCTCGACAAATTCGGCAACGA 360  
Db 301 TCTCGAATGCTTCACTCTTGAAATCATGGATAGAGCCCTCGACAAATTCGGCAACGA 360  
Qy 361 CATAGCTATTGCGCTTCACTGCTGCTGAAGATGTTGCTTTCATTTAGATGATCGAAATTTGAC 420  
Db 361 CATAGCTATTGCGCTTCACTGCTGCTGAAGATGTTGCTTTCATTTAGATGATCGAAATTTGAC 420  
Qy 421 GGGTGCACCCCTTTAGGGTTTTCAGTTTGGACACTGGGAGACTGAACCCAGAAACTTATCA 480  
Db 421 GGGTGCACCCCTTTAGGGTTTTCAGTTTGGACACTGGGAGACTGAACCCAGAAACTTATCA 480  
Qy 481 ACTTTTGTATGCGGTTCAGAGCAATATGGAATTCGCAATTCGATGATCATGTTCCCTGATGC 540  
Db 481 ACTTTTGTATGCGGTTCAGAGCAATATGGAATTCGCAATTCGATGATCATGTTCCCTGATGC 540  
Qy 541 TGTGTAGGTTTCAGGCTTGTGAGGAGTAAAGGGTTATCTCTTTCTACGAGGATGGCA 600  
Db 541 TGTGTAGGTTTCAGGCTTGTGAGGAGTAAAGGGTTATCTCTTTCTACGAGGATGGCA 600  
Qy 601 CCAAGAGTGTTCAGGGTGAGAAAGTGTAGGCCCTTTAAGGAGGGCCCTTAAGGGTCTCAG 660  
Db 601 CCAAGAGTGTTCAGGGTGAGAAAGTGTAGGCCCTTTAAGGAGGGCCCTTAAGGGTCTCAG 660  
Qy 661 AGCATGATTAAGTGTTCAGAGGAAAGACCACTGCTGCTAGTCTGAAATACCGGT 720  
Db 661 AGCATGATTAAGTGTTCAGAGGAAAGACCACTGCTGCTAGTCTGAAATACCGGT 720  
Qy 721 TGTTCAGGTTGATCCGGCTTTTGGAGGAATGGAATGGAATTCGAAAGCTTGGTGAAGTG 780  
Db 721 TGTTCAGGTTGATCCGGCTTTTGGAGGAATGGAATGGAATTCGAAAGCTTGGTGAAGTG 780  
Qy 781 GAACCCCTGTTGCAAAATGTGAAGGCCCATGACATATGGAATCTCTTAGGACCATGAATGT 840  
Db 781 GAACCCCTGTTGCAAAATGTGAAGGCCCATGACATATGGAATCTCTTAGGACCATGAATGT 840  
Qy 841 GCCTGTGAATCTCTGATGCAAAAGGATATGTTCCATTTGGGTGTGAGCCCTGCATAG 900  
Db 841 GCCTGTGAATCTCTGATGCAAAAGGATATGTTCCATTTGGGTGTGAGCCCTGCATAG 900  
Qy 901 GCCTGTTTTTACCTGGGCAACATGAAGGAGGAGGTGTGGTGGGAGGATGCCAAAGC 960  
Db 901 GCCTGTTTTTACCTGGGCAACATGAAGGAGGAGGTGTGGTGGGAGGATGCCAAAGC 960  
Qy 961 TAAGGAATGTGGTCTTCCAAAAGGAATGTAAAGCAGACAGAAAGAGGAGGTAAATGG 1020  
Db 961 TAAGGAATGTGGTCTTCCAAAAGGAATGTAAAGCAGACAGAAAGAGGAGGTAAATGG 1020  
Qy 1021 AAATGGGCTATCCCAATCCCATGCAATGTGTATGCTACCACTGTGCTGCATTTTCAA 1080  
Db 1021 AAATGGGCTATCCCAATCCCATGCAATGTGTATGCTACCACTGTGCTGCATTTTCAA 1080  
Qy 1081 CAGCCGGAATGTAGTTAACTTACAGCAGACTGGAAATTTGAGAAATTTGGCAAAATTTGGAGGA 1140  
Db 1081 CAGCCGGAATGTAGTTAACTTACAGCAGACTGGAAATTTGAGAAATTTGGCAAAATTTGGAGGA 1140  
Qy 1141 CCGAAAGGAACCATGGCTTGTGTGCTTTATGCAACCATGGTGGCCCTACTGCCAGGCTAT 1200  
Db 1141 CCGAAAGGAACCATGGCTTGTGTGCTTTATGCAACCATGGTGGCCCTACTGCCAGGCTAT 1200  
Qy 1201 GGAGGAATCTTATGTTAGCTTAGCAGACAGTATGAGGGTCAACAGGGATGAAGGTTGG 1260  
Db 1201 GGAGGAATCTTATGTTAGCTTAGCAGACAGTATGAGGGTCAACAGGGATGAAGGTTGG 1260  
Qy 1261 AAAATTTAGACAGATGAGAACAGAAATTTGCAAAAGTGAATCTGCAATTTGGGAAG 1320  
Db 1261 AAAATTTAGACAGATGAGAACAGAAATTTGCAAAAGTGAATCTGCAATTTGGGAAG 1320  
Qy 1321 CTTCCCTACGATATTAATTTTCCAAAGCATTCGTCTCGGCCCAACAATAAAGTATCCCTC 1380  
Db 1321 CTTCCCTACGATATTAATTTTCCAAAGCATTCGTCTCGGCCCAACAATAAAGTATCCCTC 1380

Qy 1381 AGAAAAGAGAGATGTTGATTCCTTGATGGCATTTGTAATGCTTAAAGATGAGGATATCA 1440  
Db 1381 AGAAAAGAGAGATGTTGATTCCTTGATGGCATTTGTAATGCTTAAAGATGAGGATATCA 1440  
Qy 1441 GGAATTTTCTTCGTTTGGTGGTTCGCAATTCACATTTGACATACGTTACGAGGGTTCCT 1500  
Db 1441 GGAATTTTCTTCGTTTGGTGGTTCGCAATTCACATTTGACATACGTTACGAGGGTTCCT 1500  
Qy 1501 TCTTTATGCTATTTACGTTGATATACCATTCGTTTACAGATTCCTCTGTGAACCTGTTGGA 1560  
Db 1501 TCTTTATGCTATTTACGTTGATATACCATTCGTTTACAGATTCCTCTGTGAACCTGTTGGA 1560  
Qy 1561 AGTGGAAATGAGGTTTATACAAATAGATACACTCAGTTTGAATGTTTAAAAAATAA 1620  
Db 1561 AGTGGAAATGAGGTTTATACAAATAGATACACTCAGTTTGAATGTTTAAAAAATAA 1620  
Qy 1621 AAAAAAAA 1629  
Db 1621 AAAAAAAA 1629

RESULT 2  
US-10-762-049-7  
; Sequence 7, Application US/10762049  
; Publication No. US20040139492A1  
; GENERAL INFORMATION:  
; APPLICANT: Falco, Saverio Carl  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCES: BB-1167-C  
; CURRENT APPLICATION NUMBER: US/10/762,049  
; CURRENT FILING DATE: 2004-01-21  
; PRIOR APPLICATION NUMBER: US/09/720,318A  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 1998-07-14  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 7  
; LENGTH: 1629  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-10-762-049-7

Query Match 100.0%; Score 1629; DB 7; Length 1629;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1629; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCACGAGGAGAGAACCCATACACAGCTAGTTAATGGCCCTCGCTTTTCACTTTCAATTTTC 60  
Db 1 GCACGAGGAGAGAACCCATACACAGCTAGTTAATGGCCCTCGCTTTTCACTTTCAATTTTC 60  
Qy 61 GCACCAACTTCCACCTTCCCATCATCGGAACCCAAATTCGCGAAATTTGGGTCAATTTAG 120  
Db 61 GCACCAACTTCCACCTTCCCATCATCGGAACCCAAATTCGCGAAATTTGGGTCAATTTAG 120  
Qy 121 GATTTCCGAGAGGCCCATTCGAGCGCGGTAAATTTCAATTTATCTCAAGACCGAGCTT 180  
Db 121 GATTTCCGAGAGGCCCATTCGAGCGCGGTAAATTTCAATTTATCTCAAGACCGAGCTT 180  
Qy 181 GGTAAAGCCGTTTAAACCCGAAACCTCCACGCAAGGATTTCCATTTCTCTCGCAGCAAC 240  
Db 181 GGTAAAGCCGTTTAAACCCGAAACCTCCACGCAAGGATTTCCATTTCTCTCGCAGCAAC 240  
Qy 241 AACCATCGTTGCTTCTGCTTCTGAGAGAAAGAGAGATTTTGAACAGATAGCCAGTGA 300  
Db 241 AACCATCGTTGCTTCTGCTTCTGAGAGAAAGAGAGATTTTGAACAGATAGCCAGTGA 300  
Qy 301 TCTCGAATGCTTCACTCTTGAATCATGATAGAGCCCTCGACAAATTCGGCAACGA 360  
Db 301 TCTCGAATGCTTCACTCTTGAATCATGATAGAGCCCTCGACAAATTCGGCAACGA 360  
Qy 361 CATAGCTATTGCGCTTCACTGCTGCTGAAGATGTTGCTTTCATTTAGATGATCGAAATTTGAC 420



Db 496 CTGAATCAGAACTACAAATTTTGTATGCGGTTGAGAAGCATATGGAATTCACATT 555  
Qy 521 GAGTACATGTTCCCTGATGCTGTGTGAGGTTGAGGTTGAGGATGTAAGGGTTATTC 580  
Db 556 GAGTACATGTTCCCTGATGCGGTTGAGGTTGAGGTTGAGGATGTAAGGTTGAGGTTCTTC 615  
Qy 581 TCCTTTACGAGGATGGCCACCAAGAGTGTGCAAGGTTGAGGTTGAGGATGTAAGGTTAAGG 640  
Db 616 TCATTTTACGAGGATGGCCATCAAGAGTGTGCGAGTAAGGTAAGGTTGAGGTTGAGG 675  
Qy 641 AGGCGCTTAAAGGCTCTCAGAGCATGGAATACCTGTGAGAGGTAAGGTAAGGTTGAGG 700  
Db 676 AGAGCCCTTAAAGGCTCTCAGAGCATGGAATACCTGTGAGAGGTAAGGTAAGGTTGAGG 735  
Qy 701 ACTAGGCTGGAATACCGGTTGTTGAGGTTGATCCGCTTTGAGGTAAGGTAAGGTTGAGG 760  
Db 736 ACTAGGCTGGAATACCGGTTGTTGAGGTTGATCCGCTTTGAGGTAAGGTTGAGGTTGAGG 795  
Qy 761 ATTGGAAGCTTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 820  
Db 796 ATTGGAAGCTTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 855  
Qy 821 TTCCCTTAGGACCATGAATGTCCTGTGAAATTCCTTGCATGCAAAAGGATATGTTTCCATT 880  
Db 856 TTCCCTTAGGACCATGAATGTCCTGTGAAATTCCTTGCATGCAAAAGGATATGTTTCCATT 915  
Qy 881 GGGTGTGAGCCCTGCATCAGGCTGTTTAACTCGGCAACATGAAGGTAAGGTAAGGTTG 940  
Db 916 GGCTGTGAGCCGTGCACAAAGCCAGTTTAACTCGGCTCAACATGAAGTAAGGTAAGGTTG 975  
Qy 941 TGGTGGAGGATGCCAAGCTTAAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGG 1000  
Db 976 TGGTGGAGGATGCCAAGCTTAAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGG 1035  
Qy 1001 AAAGAGGAGGATGTAATGGAATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATG 1060  
Db 1036 GATGCTGCCAGCTTAATGGAATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATG 1086  
Qy 1061 ACTGTGCTGACATTTTCAACAGCCGCAATGATGTTAACTTTGAGCAGGATGAGGATGAG 1120  
Db 1087 ACTGTGCTGACATTTTCAACAGCCGCAATGATGTTAACTTTGAGCAGGATGAGGATGAG 1146  
Qy 1121 AATTGGCAAAATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGG 1180  
Db 1147 AATTGGCAAAATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGG 1206  
Qy 1181 TGCCCTACTGCCAGGCTATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGG 1240  
Db 1207 TGCCCTACTGCCAGGCTATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGG 1266  
Qy 1241 TCAAAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 1300  
Db 1267 T---CAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATG 1323  
Qy 1301 AGTGAATGCAATTTGGGAGCTTCCTGAGGATGAGGATGAGGATGAGGATGAGGATGAGG 1360  
Db 1324 ACTGAATGCAATTTGGGAGCTTCCTGAGGATGAGGATGAGGATGAGGATGAGGATGAGG 1380  
Qy 1361 CCACCAATTAAGTACCTTCAGAAAGAGAGATGTTGATTCCTGATGGCATTTGTAAT 1420  
Db 1381 CAACCAATTAAGTACCTTCAGAAAGAGAGATGTTGATTCCTGATGGCATTTGTAAT 1440  
Qy 1421 GCCTTAAGATGAGGAT 1436  
Db 1441 GCTTTACGTTGATGTT 1456

RESULT 4  
US-10-424-599-134742  
; Sequence 134742, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Chou Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 134742  
; LENGTH: 2774  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_92682C.1  
US-10-424-599-134742

Query Match 56.3%; Score 916.8; DB 7; Length 2774;  
Best Local Similarity 80.7%; Pred. No. 1.2e-263;  
Matches 1126; Conservative 0; Mismatches 252; Indels 18; Gaps 4;  
Qy 41 GCTTTCACTTCTTCAATTTCCGACCAACTTCCACCTTCCCATCATCGGAACCAAACTT 100  
Db 145 GCTGCTCAGCTTCTAGCTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 204  
Qy 101 CCGCAAAATTTGGTCAATTAGGATTTCCGAGAGGCCCATTTGGAGGCGCGTTAATTTCAAT 160  
Db 205 CCTCAAAATCGGTTTCGGTTTCCGAGAGGTCGCTTGTTCGTCGTCTGTGTGTCAAT 264  
Qy 161 TTATCTCAAAAGACGGAGCTTGGTAAAGCCGTTAAACCCGAACTTCCACGCAAGGATTC 220  
Db 265 GTAACTCAACGACCTCTTGGTGAAGCCACTCAACCGCGAACCGCAACGATTC 324  
Qy 221 ATTGTTCTCTCCGAGCAACCAACCATCGTTGCTTCTCTCTCTCTCTCTCTCTCTCTCT 280  
Db 325 ATTGTTCT 381  
Qy 281 TTTGAAAGATAGCCAGTGTCTGCAATGCTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 340  
Db 382 TTTGAGCAAAATAGCGAAGACCTTTGAAATGATCTCTCTCTCTCTCTCTCTCTCTCTCT 441  
Qy 341 CTCAACAAATTTGGCAACGACATAGCTATTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 400  
Db 442 CTCGAGAAATTTGGCAACGACATAGCTATTGCTCTCTCTCTCTCTCTCTCTCTCTCT 501  
Qy 401 ATTGAGTATCGAAATTTGACGGGTCGACCTTTTGGGTTTTCAGTTTGGGACATCTGGGAGA 460  
Db 502 ATTGAGTATCGAAATTTGACGGGTCGACCTTTTGGGTTTTCAGTTTGGGACATCTGGGAGA 561  
Qy 461 CTGAACCCAGAACTTATCACTTTTGTATGCGGTTGAGAGGATTTAGGAAATTCGCATT 520  
Db 562 CTGAATCCAGAACTTATCACTTTTGTATGCGGTTGAGAGGATTTAGGAAATTCGCATT 621  
Qy 521 GAGTACATGTTTCCCTGATGCTGTTGAGGTTCAAGGATTTGAGGATTTAGGAGGTTATTC 580  
Db 622 GAGTACATGTTTCCCTGATGCTGTTGAGGTTCAAGGATTTAGTAACTTAAGGAGGCTCTTC 681  
Qy 581 TCTTTTCTACGAGGATGGGCACCAAGAGTGTGTCAGGGTTGAGAAAGGTTGAGGCTTTAAGG 640  
Db 682 TCATTTTACGAGGATGGGCATCAAGAGTGTGTCAGGATTTAGGAGGTTGAGGCTTTGAGG 741  
Qy 641 AGGCGCTTAAAGGCTCTCAGAGCATGGAATACCTGTGAGAGGTAAGGATTCAGCTCTGTT 700  
Db 742 AGAGCCCTTAAAGGCTCTCAAGAGTGTGTCAGGATTTAGGAGGATTCAGCTCTCTGTT 801  
Qy 701 ACTAGGCTGGAATACCGGTTGTTGAGGTTGATCCGCTTTTGGAGGATGATGATGAGGAGA 760  
Db 802 ACTAGGCTGGAATACCGGTTGTTGAGGTTGATCCGTTGTTGAGGAGTGTGATGAGGAGA 861  
Qy 761 ATTGGAAGCTTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 820  
Db 862 ATTGGAAGCTTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 921







APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 135400  
LENGTH: 711  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(711)  
OTHER INFORMATION: unsure at all n locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_93276C.1  
US-10-424-599-135400

Query Match 42.1%; Score 685.6; DB 7; Length 711;  
Best Local Similarity 99.1%; Pred. No. 1.5e-194;  
Matches 699; Conservative 0; Mismatches 5; Indels 1; Gaps 1;  
QY 8 GAGAGAACCCATACAGCTAGTTAATGGCCCTCGCTTTCACCTTCAATTTCCGACCA 67  
DB 7 GAGAGAACCCATACAGCTAGTTAATGGCCCTCGCTTTCACCTTCAATTTCCGACCA 66  
QY 68 ACTTCACCTTCCATCATCGGAACCCAACTTCGGAATTTGGTCAATTAGATTTGCG 127  
DB 67 ACTTCACCTTCCATCATCGGAACCCAACTTCGGAATTTGGTCAATTTAGATTTGCG 126  
QY 128 GAGAGGCCATTCGAGCGCGCTTAATTTCAATTTATCTCAAGAGCGGCTTGTAAAG 187  
DB 127 GAGAGGCCATTCGAGCGCGCTTAATTTCAATTTATCTCAAGAGCGGCTTGTAAAG 186  
QY 188 CCGTTTAAAGCCGAACCTTCAGCAAGGATTTCCATTTGTTCTCTCGAGCAACCAATC 247  
DB 187 CCGTTTAAAGCCGAACCTTCAGCAAGGATTTCCATTTGTTCTCTCGAGCAACCAATC 246  
QY 248 GTTGCTTCTGCTCTGAGAGCAAGAGGAATTTTGAACAGATAGGAGTGTCTGAC 307  
DB 247 GTTGCTTCTGCTCTGAGAGCAAGAGGAATTTTGAACAGATAGGAGTGTCTGAC 306  
QY 308 AATGCTTTCACCTCTGAATTCATGATAGAGCCCTCGACAAATTCGCAACGACATAGCT 367  
DB 307 AATGCTTTCACCTCTGAATTCATGATAGAGCCCTCGACAAATTCGCAACGACATAGCT 366  
QY 368 ATTGCTTTCAGTGGTGTCTGAAGATGTTGCTTTGATTTGATGATGCGAAATTTGAGGTCGA 427  
DB 367 ATTGCTTTCAGTGGTGTCTGAAGATGTTGCTTTGATTTGATGATGCGAAATTTGAGGTCGA 426  
QY 428 CCCTTTAGGTTTTTCAGTTTGGACACCTGGAGACTGAACCCGAAACTTATCAACTTTT 487  
DB 427 CCCTTTAGGTTTTTCAGTTTGGACACCTGGAGACTGAACCCGAAACTTATCAACTTTT 486  
QY 488 GATGCGGTTGAGAGCAATTTAGAAATTCGATTTGATGATGTTCCCTGATGCTGTTGAG 547  
DB 487 GATGCGGTTGAGAGCAATTTAGAAATTCGATTTGATGATGTTCCCTGATGCTGTTGAG 546  
QY 548 GTTTCAGGCAATTTGATGAGGATTAAGGGTTATTTCTTTTTCAGGATGAGGACCAAGAG 607  
DB 547 GTTTCAGGCAATTTGATGAGGATTAAGGGTTATTTCTTTTTCAGGATGAGGACCAAGAG 606  
QY 608 TGTTCAGGTTGAGAAAGGTAGGCTTTTAAAGAGGGCCCTTAAAGGCTCTCAGAGCATGG 667  
DB 607 TGTTCAGGTTGAGAAAGGTAGGCTTTTAAAGAGGGCCCTTAAAGGCTCTCAGAGCATGG 666  
QY 668 ATAACT -GGTCAGAGGAAGACAGTCACTGGTACTAGTCTGA 711  
DB 667 ATAACTGGGGTCAGAGGAAGACAGTCACTGGTACTAGTCTGA 711

RESULT 8

US-09-938-842A-498  
Sequence 498, Application US/09938842A  
Patent No. US20020160378A1  
GENERAL INFORMATION:  
APPLICANT: Harper, Jeff  
APPLICANT: Krepe, Joel  
APPLICANT: Wang, Xun  
APPLICANT: Zhu, Tong  
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
TITLE OF INVENTION: SAME, AND METHODS OF USE  
FILE REFERENCE: SCRIPI300-3  
CURRENT APPLICATION NUMBER: US/09/938,842A  
CURRENT FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: US 60/227,866  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: US 60/264,647  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/300,111  
PRIOR FILING DATE: 2001-06-22  
NUMBER OF SEQ ID NOS: 5379  
SEQ ID NO 498  
LENGTH: 1398  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-938-842A-498

Query Match 40.0%; Score 651.6; DB 3; Length 1398;  
Best Local Similarity 70.9%; Pred. No. 3.6e-184;  
Matches 966; Conservative 0; Mismatches 354; Indels 42; Gaps 6;  
QY 83 TCATCGGAACCCAACTTCGCAAAATTTGGTCAATTTAGGATTTTCGAGAGGCCCATTTGGA 142  
DB 67 TCATTCGAGCCAAAGTTTCGCAAAATTTGGTTCGTGAGGTTATGATCGTGTTCATGTT 126  
QY 143 GCGCGCGTTAATTTCAATTTATCTCAAGAAGCGG---AGCTTGTGAAGCCCGTTAAACGCC 199  
DB 127 GCTCTCTGTCTCTGAATCTATCTGGGAAGCGATCATCTCTGTTAAACCTTTTAAACGCT 186  
QY 200 GAACCTCCAGCGAAGGATTCATTTGTTCTCTCGAGCAACCAACCATCTGTTCTGCT 259  
DB 187 GAACCAAGACCAAGGATTCATGATTTCTCTTTCGCGCAACCAATGGTAGCAGAAATTTGCA 246  
QY 260 TCTGAGACGAAAGAG-----GAAGATTTTGAACAGATGACCCAGTGTCTCGACAAT 310  
DB 247 GAGGAAGTTGAAGTGGTTGAGATTTGAGGATTTTGAAGAGCTTCTAAGAAAGTTAGAAAT 306  
QY 311 GCTTCACTCTTTGAATTCATGATAGAGCCCTCGACAAATTTGCGCAACGACATAGCTATT 370  
DB 307 GCTTCACTCTTTGAGATTTAGCAAAAGCTCTTGAGAAATACGGAACGATATCGCAATT 366  
QY 371 GCTTCAGTGGTCTGCAAGATGTTGTTGATTTGATGATGCGAAATTTGACGGTTCGACCC 430  
DB 367 GCATTTAGTGGTGCAGAGATGTTGCTTTATTTAGTACGCTCATTTGACTGGGAGGCCA 426  
QY 431 TTTAGGTTTTTCAGTTTGGACACTGGGACCTGAACCCAGAACTTATCAACTTTTGTAT 490  
DB 427 TTTAGGATTTTAGTTTGGATACAGGAGGTTGAATCTCTGAGACGATTCGGTTTTTCGAT 486  
QY 491 GCGGTTGAGAAGCATTTATGAAATTCGCATTTGATGATACATGTTCCCTGATGCTGTTGAGGTT 550  
DB 487 GCGGTCGAGAACCATCTATGGAATTTAGGATTTAGTATATGTTTCTGATTTCTGTTGAGTT 546  
QY 551 CAGGCATTTGTCAGGAGTAAGGGTTTATTTCTTTCTTCTAGAGGATGGGCAACCAAGATGT 610  
DB 547 CAAGGTTTGGTTAGGAGCAAGGATTTGTTCTCTCTTTTATGAGGATGGTTCATCAGGATGT 606  
QY 611 TSCAGGTCGAGAAAGGTGAGGCTTTTAAAGAGGGCCCTTAAAGGCTCTCAGAGCATGATA 670  
DB 607 TGCCTGTTTCGAAAGGTGAGACCTTTGAGGCGTCTCTCAAGGCGTTTAAAGGCTTGGATT 666  
QY 671 ACTGCTCAGAGGAAGACCATGTCACCTGGTACTAGTCTGAAATACCGGTTGTTTCAGGTT 730



Db 667 ACTGGTCAGAGGAAGATCAATCTCCGGGACAAAGGCTCGAGATCCGGTGTTCAGGTT 726  
Qy 731 GATCCGCTTTTGAAGGAAATGGATGGTGAATTTGGAAGCTTGGTGAAGTGGAAACCCCTGTT 790  
Db 727 GATCCGCTTTTGAAGGTTGGATGGTGGATTTGGTAGTTGGTGAAGTGGAAATCCGGTT 786  
Qy 791 GCAAAATGTAAGGCCATGACATATGGAATCTTCCTTAGGACCATGAATGTGCCCTGTGAAT 850  
Db 787 GCGAATCTTGAAGGGAATGATGTTTGGAACTCTCTTGAGGACTATGATGTTCCGGTTAAC 846  
Qy 851 TCCTTGATCATCAAAAGGATATGTTTCCATTGGGTGTCAGCCCTGACCTAGGCTGTTTTA 910  
Db 847 ACATTGTCATGGCAGGGTATATATCGATTGGATGTGAGCCTTCGACGAAGCGTTTTA 906  
Qy 911 CCTGGGCAACATGAAAGGAGGAGGTGGTGGTGGAGGATGCAAAAGCTTAAGGAATGT 970  
Db 907 CCGGTCAGCAGAGAGAGAGGAGATGGTGGTGGGAAGATGCTAAAGCCAGGAATGT 966  
Qy 971 GGTCTTCAAAAGGAATGTAAAGCAGCAGAGAGAGGAGATGTTAATGGAATGGGCTA 1030  
Db 967 GGACTTCAAAAGGGAATGTCA---AAGAAACTCCGATGATGCTTAAAGTGAACCGGGAA 1023  
Qy 1031 TCCCAATCCCAATGCAATGGTGATGCTACCACTGTGCTGACATTTTCAACAGCCCGAAT 1090  
Db 1024 TCGAAAT-----CCGCTGTGCAGATATCTTTAAGATGAGAT 1062  
Qy 1091 GTAGTTAACTTGACGAGACTGGAATTTGAGAATTTGGCAAAATTTGGAGACCGGAAGAA 1150  
Db 1063 CTTGTGACTTTGACGAGCAGGGATTTGAGAAATTTGATGAATTTGGAGAACCGTAAAGAG 1122  
Qy 1151 CCATGGCTTGTGCTTTATGACCATGTTGGTGGCCCTACTGCCAGGCTATGGAGGAATCT 1210  
Db 1123 CTTTGGATCGTGTGCTTTATGCTCCGTGGTGGCCCTTTTGTCAAGCCATGGAAGCATCG 1182  
Qy 1211 TATGTTGACTTAGCAGACAGTTAGCAGGTCACAGGATCAACAGGATGAAGTTTGAATTTTGA 1270  
Db 1183 TATGATGAATCGGCGGATAAATGGCTGG---AAGTGGGATTAAGTTTGCCTTCCAAATTCGA 1239  
Qy 1271 GCAGATGAGAACAGAAATTTGCAAGAGTGAACCTGCAATTTGGGAAGCTTCCCTACG 1330  
Db 1240 GCAGATGGTGACAGAGGAGTTTGCTAAGCAGGAATTCAGCTCGGTAGCTTCCCTACC 1299  
Qy 1331 ATATTATTTTCCAAAGCATTGCTTCGGCCCAACATATAAGATATCCCTCAGAAAAGAGA 1390  
Db 1300 ATTCTGGTTTCCCTTAAGAACTCATC---GAGACCGATCAAGTATCCGTCTGAGAAGAGA 1356  
Qy 1391 GATGTTGATTCCTTGATGGCATTTGTAAATGCCTTAAGATGA 1432  
Db 1357 GATGTTGAGTCTTTGACTTCTTGAAATCTTGTCCGATAA 1398

RESULT 9  
US-09-938-842A-498  
; Sequence 498, Application US/09938842A  
; Publication No. US20040009476A9  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; TITLE OF INVENTION: SAME, AND METHODS OF USE  
; FILE REFERENCE: SCRI300-3  
; CURRENT APPLICATION NUMBER: US/09/938, 842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379

; SEQ ID NO 498  
; LENGTH: 1398  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-498  
Query Match 40.0%; Score 651.6; DB 3; Length 1398;  
Best Local Similarity 70.9%; Pred. No. 3 6e-184; Indels 42; Gaps 6;  
Matches 966; Conservative 0; Mismatches 354;  
Qy 83 TCATCGGAACCCAAACTTCGCAAAATTTGGTCAATTAGGATTTTCGGAGAGCCCAATTTGA 142  
Db 67 TCATTCGAGCCAAAGATTTTCGCAAAATTTGGTTCGTTGAGGTTAATTTGATCGTGTTCATGTT 126  
Qy 143 GGGCCCGTTAATTTCAATTTATCTCAAGACGG---AGCTTGGTAAAGCCCGTTAAAGCC 199  
Db 127 GCTCCTGTGTCTCTGAATCTATCTGGGAAGCGATCATCATCTGTTTAAACCTTTAAAGCGCT 186  
Qy 200 GAACCTCCAGCGCAAGGATTCATTTGCTCTCGCAGCAACCAACATCGTTTCTTCTGCT 259  
Db 187 GAACCAAGACAAGGATTCATGATTTCTCTTGGCGCAACATGGTAGCAGAAATTTGCA 246  
Qy 260 TCTGAGACGAAAGAG-----GAAGATTTTGAACAGATAGCAGTGTCTCGACAAT 310  
Db 247 GAGGAAGTTGAAGTGGTTGAGATTTGAGATTTTGAAGAGCTTGTCTAAGAAATTTAGAGAAT 306  
Qy 311 GCTTCACCTCTTGAATCATGATAGCCCTCGACAAATTCGGCAACGACATAGCTATT 370  
Db 307 GCTTCACCTCTTGAATTTAGCAAAAGCTCTTGAGAAATACCGGAAACGATATCGCCAT 366  
Qy 371 GCCTTCAGTGTGCTGAAGATTTGCTTTGATTCAGTATGCGAAATTTGAOCGGTTCGACCC 430  
Db 367 GCATTTAGTGTGCAAGAGATTTGCTCTTATTGAGTACGCTCATTTGACTGGGAGGCCA 426  
Qy 431 TTTAGGGTTTTCAAGTTTGGACACTGGGAGACTGAACCCAGAAATTTATCAACTTTTGTAT 490  
Db 427 TTTAGAGTATTTAGTTTGGATACAGGGAGGTTGAATCCTGAGAGCTATCGGTTTTCGAT 486  
Qy 491 GCGGTTGAGAAGCATTTATGGAATTCGCATTTGAGTACATGTTCCCTGATGCTGTGTGAGGTT 550  
Db 487 GCGGTGAGAAGCACTATGGGATTTAGGATTTAGTATATGTTTCTGATCTCTGTTGAGGTT 546  
Qy 551 CAGCATTTGTTGAGGAGTAAAGGGTATTTCTCTTTCTACAGAGATGGGCAACCAAGAGTGT 610  
Db 547 CAAGTTTGGTTAGGACAGAGGATTTGTTCTCTTTTATGAGGATGGTTCATCAGAGTGT 606  
Qy 611 TGCAAGGTTGAGAAAGGTGAGGCCCTTTAAAGGAGGCCCTTTAAGGGTCTCAGAGCATGATA 670  
Db 607 TGCCGTGTTTCGAAAGGTGAGACCTTTTGAGGCGTCTCTCAAGGGTTTAAAGGCTTGGATT 666  
Qy 671 ACTGTCAGAGGAAAGACCATGCTCACCTGGTACTAGGTCGAAATACCGGTTGTTTCAGGTT 730  
Db 667 ACTGTCAGAGGAAAGATCAATCTCCGGGCAAGAGTCTGAGATTCGGGTTGTTTCAGGTT 726  
Qy 731 GATCCGCTTTTGAAGGCAATGGAATTTGGAAGCTTTGGAAGCTTTGGTCAAGTGGAAACCCCTGTT 790  
Db 727 GATCCGCTTTTGAAGGTTTGAAGTGGTGGATTTGGTAGTTTGGTGAAGTGGAAATCCCGTT 786  
Qy 791 GCAAAATGTGAAGGCCCATGACATATGGAATCTTCTTTAGGACCATGAATGTGCGCTGTGAAT 850  
Db 787 GCGAATGTTGAAGGGAATGATGTTTGGAACTTCTCTGAGGACTATGATGTTTCCGGTTAAC 846  
Qy 851 TCCTTGCATGCAAAAGGATATGTTTCCATTTGGGTTGAGCCCTGACCTAGGCTGTTTTTA 910  
Db 847 ACATTGTCATGGCAGGGTATATATCGAATTCGATTCGATGAGCCTTCGACGAAGCGTTTTTA 906  
Qy 911 CCTGGCAACATGAAAGGAGGAGGTGGTGGTGGGAGGATGCGCAAGCTTAAGGAATGT 970  
Db 907 CCGGTCAGCAGAGAGAGAGGAGATGGTGGTGGGAGAGATGCTTAAAGCCAGGAATGT 966  
Qy 971 GGTCTTCAAAAGGAATTTAAGACGACGAGAAAGAGGAGGATGTTTAAATGGAATGGGCTA 1030  
Db 967 GGACTTCAAAAGGGAATGTCA---AAGAAACTCCGATGATGCTTAAAGTGAACCGGGAA 1023



Db 300 CATGGATAAGACTCTTGAGAGATTCGAGACCAATCGCAATTCCTTTAGTGGAGCTGA 359  
QY 388 AGATGTTGCTTTGATGAGTATGCGAAATGACGGGTCGACCCCTTTAGGGTTTTTCAGTTTT 447  
Db 360 AGATGTTGCAATGATTAATGATGACCTTTAACTGGAAAGCCATTTAGGGTTTTTAGTTTT 419  
QY 448 GGACACTGGGAGACTGAACCCAGAACTTATCAACTTTTGTATGCGGTTGAGAGCAATTA 507  
Db 420 AGATACAGGAGATTAACCCCTGAACGTCAGGCTCTTTGACGCACTGAGAGACAGTA 479  
QY 508 TGGAAATTCGATTCAGTACATGTTCCCTGATGCTGTTTGAAGTTTCAGGCATTTGGTGAGAG 567  
Db 480 CGGATTCGAATTCAGTACATGTTCTCTGATGCAAGTTGAGGTTCAAGCTTTAGTGAGAA 539  
QY 568 TAAGGGGTTATTCCTTTTACGAGGATGGGCACCAAGAGTGTTCGAGGTTGAGAAAGGT 627  
Db 540 CAAGGGTTTCTCTCATTCATGAAGATGGTTCATCAAGAGTGTTCGCGTGTGAGAAAGT 599  
QY 628 GAGGCTTTTAAGGAGGCGCTTAAGGCTCTCAGAGCATGATTAAGTGTTCAGAGGAAAGA 687  
Db 600 TAGACCTTTGCGTCTCTTAAGGCTCTTAAGGCTCTTAAGGCTCTTAAGGCTCTTAAGG 659  
QY 688 CCAGTCACTGGTACTAGTCTGAATACCGGTTGTTTCAAGTTGATCCGGCTTTTGAGGG 747  
Db 660 CCAATCTCCGGTACGAGATCTGATCCCTATTTGTTTCAGGTTGATCCAGTGTGTTGAAG 719  
QY 748 AATGATGTTGGAATTCGAAGCTTTGGTGAAGTGAACCCCTGTTTGCATAATGTGAAGGCCA 807  
Db 720 GTTAGATGGCGTGTGGAAGTCTGTGAAGTGAATCCCTTTGGCTAATGTGTAAGGAGC 779  
QY 808 TGACATATGNACTTCCTTAGGACCATGAATGTCCTGTGTAATTCCTTGCATGCAAGAG 867  
Db 780 TGATGTGTGAACCTTTCTGAGAACTATGGAATGTTCCGGTGAATGATTCACGACGCAAGG 839  
QY 868 ATATGTTTCCATTCGGTGTGAGCCCTGCACCTAGGACCTGTTTACCTGGGCAACATGAAG 927  
Db 840 GTATGTGTCAATCCGGTGTGAGCCGTGTACTAGGCCGGTGTCTCCAGGCCAATGAGAG 899  
QY 928 GGAAGGAGGTGTGTGGGAGGATGCAAGCTTAAGGAATGTGGTCTTCAAGAGAA 987  
Db 900 AGAAGGAAGGTGTGTGGGAGGATGCTAAGAGTAAAGAAATGTGGTCTACACAAAGGAA 959  
QY 988 TGTAAAGCAGAGAAAGGAGGATGTTAATGGAATGGGCTATCCCAATCCCATGCAAA 1047  
Db 960 CATCAAGGAGGAAGTGGTG-----CTGCAGA 986  
QY 1048 TGGTGATGCTACCACTGTGCTGTACATTTTCAACAGCCCGAATGTAGTTAACTTTGAGCAG 1107  
Db 987 CTCAAGCCCTGCTGTGCAAGAGATATTTGAAAGCAACATGTGGTTGCAATTGACAA 1046  
QY 1108 GACTGGAATTCAGAAATTTGGCAAAATTTGGAGGACCGAAAGGAACATGGCTTTGTGTGCT 1167  
Db 1047 AGGAGGGGTTGAGAATCTTTTGAAGCTAGAGAACCGTAAAGAGGCGTGGTTGGTCTACT 1106  
QY 1168 TTATGCAACATGTTGCCCTACTGCCAGGCTATGGAGGAATCTTATGTTGACTTTAGCAGA 1227  
Db 1107 TTACGCTCTTTGGTGGCCCTTTCTGCCAGGCGATGGAAGCATCGTACATCGAAATTTGGCTGA 1166  
QY 1228 CAAGTTAGCGGGTCAACAGGAGTGAAGTTTGGAAATTTAGAGCAGATGAGAACAGAA 1287  
Db 1167 GAAACTTGGCGG---AAGAGAGTTAAGTGGCGAATTTCCGAGCTGACGCTGAGCAGAA 1223  
QY 1288 AGAATTTGCAAGAGTGAACCTGCAATTTGGGAAGCTTCCCTACGATATTAATTTTCCCAA 1347  
Db 1224 GGAGTTTCTAAGCAAGAGCTTCAGTTAGGGAGCTTCCCGACGATCTCTCTTTCCGAA 1283  
QY 1348 GCATTCGTCTCGGCCAACAAATTAAGTATCCCTCAGAAAGAGATGTTGATTCCTTGAT 1407  
Db 1284 AAGAGCTCCAGG---GCTATTAAGTACCTTCCCTTCAGAGCATGAGATGTTGATTCCTCAT 1407  
QY 1408 GGCATTTTGAATGCTTTAAGATGA 1432

Db 1341 GTCGTTTGTGAATCTTCTTCGGTGA 1365

RESULT 12

US-09-938-842A-2305  
; Sequence 2305, Application US/09938842A  
; Publication No. US20040009476A9

GENERAL INFORMATION:

; APPLICANT: Harper, Jeff  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SAME, AND METHODS OF USE  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 2305  
; LENGTH: 1365  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-2305

Query Match 35.4%; Score 577.4; DB 3; Length 1365;  
Best Local Similarity 68.9%; Pred. No. 7e-162;  
Matches 886; Conservative 0; Mismatches 351; Indels 48; Gaps 5;

QY 157 CAATTATCTCAAGACGAGCTTTGTAAGCCCGTTAAAGCCGA---ACCTCAGCA 213  
Db 120 CCATTTCTCAGAGAGCTTACTATGAAACCTTTAAACGCTGAGTCACATTCACGAAG 179  
QY 214 GGATTCATTTCTCTCGCAGCAACACATCGTTGCTTCTGCTCTGAGA-----C 267  
Db 180 CGAATCTTGGGTTACTCGTCTTCACTCTAATTCCTTGAAGTTGAGAGAAAGGAGG 239  
QY 268 GAAAGAGGAAGATTTTGAACAGATAGCCAGTGTCTCGAACAATCTTCACTCTTGAAT 327  
Db 240 AGAAGTTGAAGACTTTCAGCAACTTGTCTAAAAGCTTGAAGATGCTTCTCCACTTGAAT 299  
QY 328 CATGGATAGAGCCCTCAGCAAAATTCGGCAAGCATAGCTATTGCTTCAGTTCAGTGTGCTGA 387  
Db 300 CATGGATAAAGCTCTTGAGAGATTCGGAGACCAAAATCGCAATTCCTTTAGTGGAGCTGA 359  
QY 388 AGATGTTGCTTTGATTCAGTATGCAAAATTCAGCGGTCGACCTTTAGGGTTTTTCAGTTT 447  
Db 360 AGATGTTGCAATTTGAATATGCAACGTTTAACTGGAAGCAATTTAGGGTTTTTAGTTT 419  
QY 448 GGACACTGGGAGACTGAACCCAGAAACTTATCAACTTTTGTATGCGGTTGAGAAAGCAATTA 507  
Db 420 AGATACAGGAGATTAACCCCTGAAACGTCAGGCTCTTTGACGAGTCGAGAACAGTGA 479  
QY 508 TGGAAATTCGATTCAGTACATGTTCCCTGATGCTGTTGAGGTTTCAGGCATTCAGGATTCAGGAG 567  
Db 480 CGGATTCGAATTCAGTACATGTTTCCCTGATGCTGAGTTGAGGTTTCAAGCTTTAGTGAGAA 539  
QY 568 TAAGGGGTTATTCCTTTTACAGGAGTGGGCACCAAGAGTGTTCAGGGTTGAGAAAGGT 627  
Db 540 CAAGGGTTTGTTCATTCATTAAGATGGTTCATCAAGAGTGTTCGCGTGTGAGAAAGT 599  
QY 628 GAGGCTTTTAAGGAGGCGCTTTAAGGCTCTCAGAGCATGGAATACTGCTCAGAGAAAGA 687  
Db 600 TAGACCTTTGCGTCTGCTCTTTAAGGCTCTTAAAGCTTGAAGTTACAGGACAGAGAAAGA 659  
QY 688 CCAGTCACTGGTACTAGGCTCTGAATACCGGTTGTTTCAAGTTGATCCGGCTTTTGAGGG 747  
Db 660 CCAATCTCCGGGTACGAGATCTGAGATCCCTTATTTGTTTCAAGTTGATTCAGTGTGTAAGG 719

```

Qy 748 AATGGATGGTGAATTTGGAAGCTTTGGTCAAGTGAACCCCTGTTGCAAAATGTGAAGGGCCA 807
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 720 GTTAGATGGCGTGTGGAAGCTTTGTGAAGTGAATCCTTTGGCTAATGTGAAGAGC 779
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 808 TGACATATGGAATCTTCCTTAGGACCATGAATGTGCGCTGTGAATTCCTTGTGANTGCAAAAG 867
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 780 TGATGTGGAATCTTCCTGGAACATATGGAATGTGCGTGTGAATGCAATGCAAGCACAAG 839
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 868 ATATGTTTCCATTTGGGTGTGAGCCCTGCACCTAGGCTGTTTACCTGGGCAACATGAAG 927
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 840 GTATGTCTAATCGGGTGTGAGCGTGTACTAGCCCGTGTCTCCAGGCCAACATGAGAG 899
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 928 GGAAGGAGGTGTGTGGGAGGATGCCAAAGCTTAAGGAATGTGCTTCCACAAAGGAAA 987
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 900 AGAAGGAAGTGTGTGGGGAAGATGCTAAAGCTTAAGAAATGTGCTCTACACAAAGGAA 959
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 988 TGTAAAGCAGAGAAAGAGAGGAGTGTAAATGGAATGGGCTATCCCAATCCCAATGCAAA 1047
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 960 CATCAAGGAGGAAGATGGTG-----CTGCAGA 986
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1048 TGGTGATGCTACCACTGTGCTGACATTTTCAACAGCCCGAATGTAGTTAACTTTGAGCAG 1107
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 987 CTCAAGCCCTGCTGTGCAAGAGATATTGAAAGCAACAATGTGTTGCAATTGACAA 1046
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1108 GACTGGAATTTGGAATTTGGCAAAATTTGGAGGACCGAAAGGAACCATGGCTTGTGTGCT 1167
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1047 AGGAGGGTGTGAGAATCTTTTGAAGCTAGAGAACCGTAAAGAGCGTGTGTTGCTGACT 1106
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1168 TTATGACCAATGTGCGCCCTACTGCGAGGCTATGGAGGAATCTTATGTGACTTAGGAGA 1227
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1107 TTACGCTCTCTGTGTGCGCCCTTTCTGCCAGGGGATGGAAGCATCGTACATCGAAATGGCTGA 1166
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1228 CAAGTTAGCGGTCAACAGGATGGAAGTTGGAATTTAGAGCAGATCGAGAACAGAA 1287
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1167 GAACTTGGCGG---AAAGAGTTAAAGTGGCGAAATTCGAGCTGACCGTGAGCAGAA 1223
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1288 AGAATTTGCAAGAGTGAACCTGCAATTTGGGAAGCTTCCCTACGATATATTTTCCCAAA 1347
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1224 GGAGTTTGTCAAGCAAGAGCTTCAGTTAGGAGCTTCCCGACGATACTCTCTTCCGAA 1283
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1348 GCATTCGCTCGGCCAACAATAAGTATCCCTCGAAGAGAGATGTTGATTCCTTGAT 1407
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1284 AAGAGCTCCACGG---GCTATTAAGTACCCCTCAGAGCAGATAGAGATGTTGATTCATCAT 1340
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1408 GGCATTTGTAATGCCTTTAGATGA 1432
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1341 GTCGTTTGTGAATCTTCTTCGGTGA 1365
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

# RESULT 13

```

US-10-731-525-9
; Sequence 9, Application US/10731525
; Publication No. US2004012140A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/10/731,525
; PRIORITY FILING DATE: 2003-12-09
; PRIORITY FILING DATE: US/09/720,318A
; PRIORITY FILING DATE: 2000-12-21
; PRIORITY FILING DATE: 60/092,833
; PRIORITY FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 1827
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-731-525-9

```

```

Query Match 33.0%; Score 538.2; DB 7; Length 1827;
Best Local Similarity 68.1%; Pred. No. 5e-150;
Matches 801; Conservative 0; Mismatches 358; Indels 18; Gaps 3;

Qy 257 GCTTCTGAGAGAAAGGAGAGATTTTGAACACATAGATCCAGTGTCTCGACAATGCTTCA 316
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 323 GCATCTCGCGTGGCGCGCTGGAATAAGAGAGCCCTGGCGAGGAGCTGGTGGCGCGCTCG 382
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 317 CCTTTGAAATCATGTATGAGCCCTCGACAAATTCGGCAACGACATAGTATATGCTTTC 376
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 383 CCCTGGAGATCATGTATGCTGCTCGACATGTTGGCTCCGAAATCGGCATCGCCTTC 442
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 377 AGTGTGCTGAAGATGTTGCTTTGATTTAGTATGCGAAATTTGACGGTTCGACCTTTAGG 436
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 443 AGTGTGCGGAGGAGCTGGGCCCTCATCGAATACGGAATCTGATGACGCCCTTCAGG 502
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 437 GTTTTCAGTTTGGACACTGGGAGACTGAACCCAGAACTTTATCAACTTTTGTATGCCGTT 496
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 503 GTGTTTCAGCCTTGACACTGGGCGACTGAACCCAGAGACATACGAACCTCTTCGACAAGTG 562
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 497 GAGAAGCATTTGGAATTTGCAATTTGAGTACATGTTCCCTGATGCTGTTGAGGTTTCAGGCA 556
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 563 GAGAGCACTATGATATCCACATCGAGTACATGTTCCCTGAGCCAGGAGGTGCAAGAC 622
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 557 TTGCTGAGGAGTAAAGGGTTTATCTCTTTCTACAGAGATGGGCAACGAAGTGTTCAGG 616
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 623 CTTGTGAGGAGCAAGGCCCTCTCTCTTTCTACAGGAGCGACACACAGGAGTGTCTGAGG 682
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 617 GTGAGAAAGTGAAGGCTTTAAGGAGGCCCTTAAGGCTCTCAGAGCATGGATTAACCTGT 676
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 683 GTGAGAAAGTTCGGGCCCTTTGAGGAGGGGCCCTCAAGGGCCTGGAATCACCAGG 742
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 677 CAGAGAAAGACCACTGCTGCTGCTAGTCTGGAATACCGTGTTCAGGTTGATCCG 736
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 743 CAGAGAAAGATCAGTCCCTCGCAACAGACCAAGCATCTCTGTTGTTCAAGTTGATCCG 802
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 737 GCTTTTGAAGGAAATGGAATTTGGAAGCTTGGTGAAGTGAACCCCTGTTTGAACAT 796
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 803 TCTTTTGAAGGGCTGAGTGGTGGAGCGGTAGCTTGATCAAGTGAACCCCTGTGGCTAAT 862
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 797 GTGAAGGCCATCAGATATGGAATCTCTTAGACCATGAATGTGCTGTGAATTCCTTG 856
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 863 GTGGATGGCAAGGATATCTGGACCTTCTCAGAGCAATGATGTCCTGTGTAACACCTG 922
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 857 CATGCAAAAGGATATGTTTCCATTTGGTGTGAGCCCTGCACTAGGCTGTTTACCTGGG 916
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 923 CATGCTCAGGCTACGTCTCCATTTGGTGTGAGCCCTGCAAGGCGGCTGTTGCCGGGG 982
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 917 CAACATGAAGGGAAGGAGGTGTTGGTGGAGGATGCCAAAGCTAAGGAATGTGCTTT 976
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 983 CAGCAGAGAGGGAAGGAGGTGTTGGTGGAGGAGCGCCACGCGCAAGGAGTGGCGCTG 1042
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 977 CACAAAGGAATGTAAAGCAGAGAGAGGAGGATGTTAATGGAATGGGCTATCCAA 1036
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1043 CACAAAGGTA-----ACATCGACAAAGGAAGGTTCAGGCACCCAAAGGTGGCGGTC 1090
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1037 TCCCATGCAAAATGCTGATGCTACCACTGTGCTCGCTGACATTTTCAACAGCCCGAATGTAGTT 1096
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1091 AACGGCAACGGCTCGCGCGGAGGCGAGGCCCGCCAGACATCTTCAGAGCAGGCGCATCGTC 1150
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1097 AACTTGAGCAGGACTGGAATTTGAGAAATTTGGCAAAATTTGGAGGACCGAAAGAACCATGG 1156
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1151 AACCTCACCGTCCCGGATCGAGAACCTCTCTGGGCTCGAGAACCGCGCGGAGCCGCTGG 1210
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1157 CTTGTTGTGCTTTATGCACCATGTTGCCCTACTGCGAGGCTATGAGGAACTTATGTT 1216
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1211 CTCACCGTCTCTACGCTCCCTGTGTCCTATCTGCGAGGCAATGGAAGGCTCTCATGTT 1270
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1217 GACTTAGCAGACAAGTTAGCAGGCTCAACAGGATGAAGGTTGGAATAATTTAGAGCAGAT 1276
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1271 GAGCTGGCGAGAAAGTGAAGCGGCT---CAGGATCAAGGTGGCAAGTTCCGCGCGGAC 1327
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1277 GGAGAAACAGAAAGAAATTTGCAAAAGAGTGAACTGCAATTTGGGAAGCTTCCCTACGATATTA 1336
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db	1328	GGCGAGCAGAAGCCATTTCGGCGAGCGCGAGCTGCAACTACAGAGCTTCCCGAGATCCTC	1387
Qy	1337	TTTTTCCCAAGCATTTCGTTCGSCCAACAATAAAGTATCCCTCAGAAAAAGAGAGATGTT	1396
Db	1388	CTGTTCCTCCCGCGCCACCGTGAAGC-----CCATCAAGTACCCGTCCTCGAAGAGGGACGTC	1444
Qy	1397	GATTCTCTTGATGGCATTTTGTAAATGCTTAAAGATGAG	1433
Db	1445	CAGTCCCTCCTCGCTTCGTGAACAGCCTCAGATGAG	1481
RESULT 14			
US-10-762-049-9			
; Sequence 9, Application US/10762049			
; Publication No. US20040139492A1			
; GENERAL INFORMATION:			
; APPLICANT: Falco, Saverio Carl			
; APPLICANT: Allen, Stephen M.			
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins			
; FILE REFERENCE: BB-1167-C			
; CURRENT APPLICATION NUMBER: US/10762,049			
; CURRENT FILING DATE: 2004-01-21			
; PRIOR APPLICATION NUMBER: US/09/720,318A			
; PRIOR FILING DATE: 2000-12-21			
; PRIOR APPLICATION NUMBER: 60/092,833			
; PRIOR FILING DATE: 1998-07-14			
; NUMBER OF SEQ ID NOS: 12			
; SOFTWARE: Microsoft Office 97			
; SEQ ID NO 9			
; LENGTH: 1827			
; TYPE: DNA			
; ORGANISM: Triticum aestivum			
US-10-762-049-9			
Query Match 33.0%; Score 538.2; DB 7; Length 1827;			
Best Local Similarity 68.1%; Pred. No. 5e-150;			
Matches 801; Conservative 0; Mismatches 358; Indels 18; Gaps 3;			
Qy	257	GCTTCTGAGACGAAGAGGAATTTTGAACAGATAGCCAGTGTATCGACAAATGCTTCA	316
Db	323	GCATCTGCGTGGCGCGCTGGACTACGAGGCCCTTGGCGCAGGAGCTGGTGGCGCGCTCG	382
Qy	317	CCTCTTGAATTCATGGATAGAGCCCTCGACAAATTCGSCAACAGCATAGCTATTGGCTTC	376
Db	383	CCGCTGGAGATCATGGATCGTGGCTGCAATGTTGGCTCCGAATCGCCATCGCCTTC	442
Qy	377	AGTGGTGTGAAGATGTTGCTTTGATTGAGTATCGAAATTCAGGGTTCGACCCCTTTAGG	436
Db	443	AGTGGTCCGAGGACGTGGCCCTCATCGAAATCGCGAAATGACTGGAGCGCCCTTCAGG	502
Qy	437	GTTTTCAGTTTGGACATCTGGGAGACTGAAACCCAGAAACTTATCAACTTTTGTGATCGGTT	496
Db	503	GTGTTTCAGCCTTGAACATGGGCGCATGAAACCCAGAGACATACGAACTCTTCGACAAAGTG	562
Qy	497	GAGAGCATATAGGAATTCGATTCAGTACATGTTCCCTGATGCTGTTGAGGTTTCAGGCA	556
Db	563	GAGAAGCATATGTTGATTCATCGATCGAGTACATGTTTCCCTGAGGCCAGCGAGGTGCAAGAC	622
Qy	557	TTGGTGGAGGTAAGGGGTTATTCTCTTTCTACGAGGATGGGCACCAAGAGTGTTCGAGG	616
Db	623	CTTGTGAGGACNAGGGCTCTTCTCTTTCTACGAGGACGACACCCAGAGTGTTCGACG	682
Qy	617	GTGAGAAAGGTGAGGCCCTTTAAGGAGGGCCCTTAAAGGTCCTCAGAGCATATGATACTGGT	676
Db	683	GTGAGGAAGGTTTCGGCCCTTGGAGGAGGGCCCTCAAGGGCCCTCAAGGGCTGGATCACC	742
Qy	677	CAGAGGAAGACCACTGACCTGGTACTAGGTCGAAATACCGGTTGTTTCAGGTTGATCCG	736
Db	743	CAGAGGAAGGATCAGTCCCTTGGCCACAGAGCCATCCCTTGTTGTTCAAGTTGATCCG	802
Qy	737	GCTTTTGGAGGGAATGGATGTTGGAAATTCGAAAGCTTGGTGAAGTGGAAACCTCTTTGCAAT	796

Matches 774; Conservative 0; Mismatches 415; Indels 3; Gaps 1;	
Qy	243 CCATCGTTGCTTCTGCTTCTGAGACGAAAGAGAGATTTTGAACAGATAGCCAGTCATC 302
Db	303 CCNCGCGCGCCGTCGCGACGAGCGCGGTGACTAGAGGCCCTGCGCGCGAGC 362
Qy	303 TCAGCAATGCTTCACTCTTGAATCATGATAGACCCCTCGACAAATTCGGCAACGACA 362
Db	363 TTGAGGCGCGCTGCGCGCTGAGATCATGATCGGCGCTGGCCATGTTCCGGTCCGAAA 422
Qy	363 TAGCTATTGCTTTCAGTGTGCTGAGATCTTCTTGTGATGATGATGCGAAATTCACGG 422
Db	423 TCGCCATTCGCTTTCAGCGCGCGGAGAGACCTGGCGCTGATCGAGTACGCGANGCTGACGG 482
Qy	423 GTCGACCCCTTTAGGGTTTTCAGTTTGGACACTGGGAGACTGAACCCAGAAACTTATCAAC 482
Db	483 GCGCTCCCTTCGGGTGTTTCAGCTTGAACACGCGGCGGCTCAACCCGAGACGTACGAGC 542
Qy	483 TTTTGTGATGCGGTTGAGAAGCATATGGAATTCGCATGAGTACATGTTCCCTGATGCTG 542
Db	543 TCTTCGACAGGGTGGAGAGCATACGCGCATCGGCATCGAGTACATGTTCCCGAGCGGA 602
Qy	543 TTGAGGTTCAAGGATTCGTCAGGATGAGGTTATCTCTTCTTACGAGGATGGGACC 602
Db	603 GCGAGGTGAGGAGCTGGTGGCAACCAAGGCGCTTCTCTTCTACGAGGACGCGCAC 662
Qy	603 AAGAGTGTTCAGGGTGAAGAGGTGAGGCTTTTAAAGAGGGCGCTTAAAGGGTCTCAGAG 662
Db	663 AGAGTGTCTCGGCTGGCAAGGTGGCGCCCTTGGCGAGGGCGCTCAGGGGCTCAGGG 722
Qy	663 CATGGAATACTGTCTAGAGAAAGACAGTCACTCTGATCTAGGTCTGAAATACCGGTTG 722
Db	723 CGTGGATCACCGGCGAGAGAGAGACAGTCCCGCGGCAACGAGGCGCAGCATCCCCCTGG 782
Qy	723 TTCAGGTTGATCGGCTTTTGAAGGATGATGATGGTGAATGGAAGTTGGTGAAGTGA 782
Db	783 TCAGGTGCAACCTTCTCTGAAGGCTTGAACGCGGGGCTGGTGGTGGTGGTGGTGGTGG 842
Qy	783 ACCCTGTGCAAAATGTGAAGGGCATGACATATGAACTTCTTACGAGCCATGAATGTGC 842
Db	843 ACCCGTGGCCACGTTCGACGGCAAGGACATCTGGACTTCTCCGGACCATGGAGCTCC 902
Qy	843 CTGTGAATTCCTTGCATGCAAAAGGATATGTTTCCATTTGGGTGTGAGCCCTGCACCTAGGC 902
Db	903 CTGTCAACGCGCTGCACGCGCCAGGCTACGTGTCCATCGGGTGGAGCGGTGCACCGGC 962
Qy	903 CTGTTTTCCTGGGCAACATGAAGGAGGAGGTGGTGGTGGGAGGATGCCAAAGCTA 962
Db	963 CTGTCTCGCGGGGCGACGAGCGGGAAGGCGGCTGGTGGTGGGAGGACGCCAAGGCCA 1022
Qy	963 AGGAATGTGGTCTTTCACAAAGGAAATGTAAAGCAGCAGAAAGAGAGGATGTTAATGGAA 1022
Db	1023 AGAGTGGCGCTTCCACAAAGGGCAACATCGAAGGACGCGCGGCGCGCGCGCGCGCGG 1082
Qy	1023 ATGGGCTATCCCAATCCCATGCAAAATGGTATGCTACCATGTGCTGACATTTTCAACA 1082
Db	1083 CGCGCGCCCGCCAGGACTGCGCAACGCGCAACGCTCGCGCGCGCGCGCGGACATCTTCGAGA 1142
Qy	1083 GCGCGAATGTAGTTAATTGAGCAGGACTGGGAATTGAGAAATTGGCAAAATTTGGAGGACC 1142
Db	1143 GCGCGCGCGCTGCTGCTCCCTCACCGCGCGGGTTCGAGAACCTGTGCGGCTGGAGAGCC 1202
Qy	1143 GAAAGGAACCATGGCTTGTGTCTTTATGCACCATGTGGTCCCTTACTGCGAGGCTATGG 1202
Db	1203 GCGCGAGCGGTGCTGCTGCTGTGACGCGCTTGGTGGCGCTTCTGCGAGGCCATGG 1262
Qy	1203 AGGAATTTATGTTGATTTAGCAGACAAGTTAGCAGGGTCAACAGGGATGAAGGTTGGAA 1262
Db	1263 AGGCTCTTACGTGGAGCTGGCGAGAGGCTGGCGGGTCCGCGCGCTGAAGGTGGCCC 1322
Qy	1263 AATTTAGACAGATGAGAACAGAAAGATTTGCAAGAGATGAACTGCAATTTGGGAAGCT 1322
Db	1323 GGTTCGCGCGCGACGCGAGCAGAAAGCGTTTCGCGAGGCGGAGCTGCAGAGCT 1382

Search completed: February 21, 2006, 00:17:31  
Job time : 1414 secs

Qy	1323 TCCTACGATATTATTATTTTCCCAAGCATTCGTCTCGGCCCAACAATAAAGTATCCCTCAG 1382
Db	1383 TCCCCACCGTCTCTCTGTTCCCGCAGCGGCAACCGCCCGGC---CCATCAAGTACCGTCGG 1439
Qy	1383 AAAAGAGAGATGTTGATTCCTTGTATGTCATTTTGTAAATGCTTTAAGATGAGG 1434
Db	1440 AGAAGAGGAGCGTCGACTCGCTCTCGCTTCGTCAACAGCCTCCGCTGAGG 1491

**THIS PAGE LEFT BLANK**



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 21:14:31 ; Search time 540 Seconds  
(without alignments)

6403.821 Million cell updates/sec

Title: US-10-731-525-7

Perfect score: 1629

Sequence: 1 gcacgaggagagaccata.....taaaaaaaaaaaaaaaaaaa 1629

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7204323 seqs, 1061406715 residues

Total number of hits satisfying chosen parameters: 14408646

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications NA.New.\*
- 1: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
  - 2: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
  - 3: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
  - 4: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
  - 5: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
  - 6: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
  - 7: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
  - 8: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
  - 9: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*
  - 10: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq2.\*
  - 11: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq3.\*
  - 12: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq4.\*
  - 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	568	34.9	1082144	12	US-11-117-187-211
2	41.8	2.6	1161	8	US-10-750-185-59729
3	41.8	2.6	1161	8	US-10-750-623-59729
4	39.8	2.4	4751	7	US-10-511-538-86
5	39.4	2.4	1366	6	US-09-925-065A-685225
6	38.8	2.4	615	6	US-09-925-065A-827921
7	38.8	2.4	622	6	US-09-925-065A-827195
8	38.8	2.4	194553	12	US-11-098-686-8738
9	38.6	2.4	569	6	US-09-925-065A-335502
10	38	2.3	661	6	US-09-925-065A-687471
11	37.6	2.3	535	6	US-09-925-065A-159939
12	37.6	2.3	615	6	US-09-925-065A-827920
13	37.6	2.3	622	6	US-09-925-065A-852196
14	37	2.3	519	6	US-09-925-065A-466709
15	36.8	2.3	459	6	US-09-925-065A-192552
16	36.8	2.3	618	6	US-09-925-065A-423102
17	36.8	2.3	618	6	US-09-925-065A-423103
18	36.8	2.3	618	6	US-09-925-065A-423104
19	36.8	2.3	1690	12	US-11-090-439-43
20	36.6	2.2	600	12	US-11-136-527-6775

ALIGNMENTS

RESULT 1

US-11-117-187-211  
; Sequence 211, Application US/11117187  
; Publication No. US20050266560A1  
; GENERAL INFORMATION:  
; APPLICANT: PREUSS, DAPHNE  
; APPLICANT: COPENHAVEN, GREGORY  
; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS  
; FILE REFERENCE: ARCD:309US  
; CURRENT APPLICATION NUMBER: US/11/117,187  
; PRIOR FILING DATE: 2005-04-28  
; PRIOR APPLICATION NUMBER: US/09/531,120  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: 60/125,219  
; PRIOR FILING DATE: 1999-03-18  
; NUMBER OF SEQ ID NOS: 212  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 211  
; LENGTH: 1082144  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-11-117-187-211

Query Match 34.9%; Score 568; DB 12; Length 1082144;  
Best Local Similarity 73.5%; Pred. No. 2e+142;  
Matches 791; Conservative 0; Mismatches 255; Indels 30; Gaps 4;  
Qy 370 TGCCTTCAGTGGTGAAGATGCTGCTTTGATTGATGCGAAATTGACGGGTCCACC 429  
Db 354529 TGTTTCAGTGGTGAAGATGCTGCTTTGATTGATGCGTCAATTTGACTGGAGGCC 354588  
Qy 430 CTTTAGGGTTTTCAGTTCGACACTGGGAGACTGAACCCAGAACTTATCAACTTTTGA 489  
Db 354589 ATTAGATGATTAGTTTGGATACAGGAGGTTGATCTCGAGACGATCGTTTTCGA 354648  
Qy 490 TCGGTTGAGAGCATATGGAATTCGATTTAGTATGATGTCCTTCGATGCTGTGAGGT 549  
Db 354649 TCGGTTGAGAGCATATGGAATTCGATTTAGTATGATGTCCTTCGATGCTGTGAGGT 354708  
Qy 550 TCAGGCATTTGGTGGAGGATGAGGGTTATCTCTTCTACGAGGATGGGACCAAGAGTG 609  
Db 354709 TCAAGGTTTGGTTAGGAGCAAGGATGTTCTCTCTTTATGAGGATGGTTCATCAGGAGTG 354768

```
Qy 610 TTCCAGGGTGAAGAGTGGAGCCCTTTAAGAGGGCCCTTAAGGGTCTCAGAGCATGGAT 669
Db 354769 TTGCCGGTGTTCGAAGGTGAGACCTTTTGAGCGGTCTCTAAGGGTTTAAAGGCTTGGAT 354828
Qy 670 AACTGGTTCAGAGGAAAGACCAAGTCACTCGTGTACTAGGTCTGAATACCGGTGTTCAGGT 729
Db 354829 TACTGGTTCAGAGGAAAGATCAATCTCCGGGACAAAGTCTGAGATTCCGGTGTTCAGGT 354888
Qy 730 TGATCCGGCTTTTGAAGGAATGGATGGTGAATTTGAAGCTTGGTGAAGTGAACCTGT 789
Db 354889 TGATCCGGCTTTTGAAGGTTCGATGTGGATGTGGAGTTGGTAGTTTGGTGAAGTGAATCCGGT 354948
Qy 790 TGCAAAATGTGAAGGCCATGACATATGGAATCTTCTAGGACCAATGAATGCTGCTGAA 849
Db 354949 TGCAAATGTGAAGGGAATGATGTTCGAACTTCTTGAAGCACTAGGATGTCCGGTTAA 355008
Qy 850 TTCTTCGATGCAAAAGGATATGTTTCCATTTGGGTGTGAGCCCTGCACCTAGGCCTGTTTT 909
Db 355009 CACATTGATGCGCAGCGGTATATATCGATTGGATGTGAGCCTTGACGAAGCGTTTT 355068
Qy 910 ACTGGGCAACATGAAAGGGAAGGAGGTGTGTGGAGGATGCCAAAGCTAAGGAATG 969
Db 355069 ACCGGGTGAGCAGAGAGAGGAGATGTGTGGGGAAGATGCTAAAGCCAAAGGAATG 355128
Qy 970 TGGTCTTCAAAAGGAATGTAAAGCAGCAGAGAGGAGATGCTTAATGGAATGGCT 1029
Db 355129 TGGACTTCAAAAGGGAATGTCA---AAGAAAATCCGATGATGCTAAAGTGAACGGGGA 355185
Qy 1030 ATCCCAATCCCATGCAAAATGATGTGTACCACTGTGCTGCACATTTTCAACAGCCCGAA 1089
Db 355186 ATCGAAT-----CCGCTGTTGCAGATATCTTTAAGATGAGAA 355224
Qy 1090 TGTATTAACCTGAGCAGGACTGGAATGGAATTTGGCAAAATTTGAGAGCCGAAGGA 1149
Db 355225 TCTTGTGACTTTGAGCAGGAGGGAATTTGAAATTTGATGAAGTTGGAGAACCGTAAAGA 355284
Qy 1150 ACCATGGCTTGTGCTTTATGACCAATGCTGCTCCCTACTGCCAGGCTATGGAGGAATC 1209
Db 355285 GCCTTGGATGCTGCTGCTTTATGCTCGGTGTGCCCCCTTTTGTCAAGCCATGGAAGCATC 355344
Qy 1210 TTATGTTGACTTAGCAGCAAGTTAGCAGGCTCAACAGGGATGAAGTTGGAAATTTAG 1269
Db 355345 GTATGATGAACCTGGCGGATTAATTTGGCTGG---AAGTGGATTAAGTTGGCAAAATTCAG 355401
Qy 1270 AGCAGATGGAGAACAGAAAGAAATTTGAAAGATGAATGCAATTTGGAGACTTCCCTAC 1329
Db 355402 AGCAGATGCTGACCAAGAGGATTTGCTTAAGCAGGAATTCAGCTCGGTAGCTTCCCTAC 355461
Qy 1330 GATATTATTTTCCCAAGCAATTCGTCTCGGCCCAACATAAAGTATCCCTCAGAAAAGAG 1389
Db 355462 CATCTGGTTCCTTCCCTAAGAACTCATC---GAGACCGATCAAGTATCCGTCTGAGAAGAG 355518
Qy 1390 AGATGTTGATTCCTTTGATGGCATTTGTAATGCTTTAAGATCAGGATATCAGGAAA 1445
Db 355519 AGATGTTGATTCCTTTGACTTCGTTCTTGAACTTCCTCCGATGAAGTAAACCCACAAA 355574
```

## RESULT 2

```
US-10-750-185-59729/c
; Sequence 59729, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
```

```
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 59729
; LENGTH: 1161
; TYPE: DNA
; ORGANISM: Bovine 19866881617930
US-10-750-185-59729
```

```
Query Match 2.6%; Score 41.8; DB 8; Length 1161;
Best Local Similarity 56.0%; Pred. No. 0.62;
Matches 79; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
```

```
Qy 1367 ATAAAGTATCCCTCAGAAAAGAGAGATGTTGATTCCTTGATGGCATTTGTAATGCCCTTA 1426
Db 886 ATAAACATTTCTCACAACTCACACATGTTGATCACCTATTGACATTTTATGAAGGAATTG 827
Qy 1427 AGATGAGATATCAGGAAATTTTCTCGTTTTGGTGGTTCGAAATTCACATTTGACTATACG 1486
Db 826 TGTGATGAGATTTTGACTATTTCTTATCTCTATAACTTTAAATGGAATCTTGATGAAT 767
Qy 1487 TACAGCGGTTCCCTCTTTTAT 1507
Db 766 TACTGAGGTATTTTTCATTAT 746
```

## RESULT 3

```
US-10-750-623-59729/c
; Sequence 59729, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 59729
; LENGTH: 1161
; TYPE: DNA
; ORGANISM: Bovine 19866881617930
US-10-750-623-59729
```

```
Query Match 2.6%; Score 41.8; DB 8; Length 1161;
Best Local Similarity 56.0%; Pred. No. 0.62;
Matches 79; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
```

```
Qy 1367 ATAAAGTATCCCTCAGAAAAGAGAGATGTTGATTCCTTGATGGCATTTGTAATGCCCTTA 1426
Db 886 ATAAACATTTCTCACAACTCACACATGTTGATCACCTATTGACATTTTATGAAGGAATTG 827
Qy 1427 AGATGAGATATCAGGAAATTTTCTCGTTTTGGTGGTTCGAAATTCACATTTGACTATACG 1486
Db 826 TGTGATGAGATTTTGACTATTTCTTATCTCTATAACTTTAAATGGAATCTTGATGAAT 767
Qy 1487 TACAGCGGTTCCCTCTTTTAT 1507
Db 766 TACTGAGGTATTTTTCATTAT 746
```

## RESULT 4

```
US-10-511-538-86
; Sequence 86, Application US/10511538
```

```
Publication No. US20060026700A1
GENERAL INFORMATION:
APPLICANT: Origene Technologies, Inc
TITLE OF INVENTION: TISSUE SPECIFIC GENES AND GENE CLUSTERS
FILE REFERENCE: 16U 200 PCT
CURRENT APPLICATION NUMBER: US/10/511.538
PRIORITY FILING DATE: 2004-10-18
PRIORITY FILING DATE: 2002-04-16
PRIORITY FILING DATE: 2002-04-16
PRIORITY FILING DATE: 2002-09-20
PRIORITY FILING DATE: 2002-11-07
PRIORITY FILING DATE: 2002-04-24
PRIORITY FILING DATE: 2002-05-01
PRIORITY FILING DATE: 2002-05-01
PRIORITY FILING DATE: 2002-05-20
PRIORITY FILING DATE: 2002-08-16
NUMBER OF SEQ ID NOS: 344
SOFTWARE: PatentIn version 3.1
SEQ ID NO 86
LENGTH: 4751
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (382)..(1929)
OTHER INFORMATION:
US-10-511-538-86

Query Match      2.4%; Score 39.8; DB 7; Length 4751;
Best Local Similarity 48.9%; Pred. No. 4.3; Indels 0; Gaps 0;
Matches 107; Conservative 0; Mismatches 112;

QY 1411 ATTGTAAATGCCCTTAAGATGAGGATATCAGGAAATTTCTTCGTTTGGGTTGCAATT 1470
DB 4533 ACTTTGGAATTTATCTGAGTAGAATAATCAGAAGACATTATCTAATCTTTGTAGATACACTGT 4592
QY 1471 CCACCTTGACATACGTACAGCGGTCCTCTTTATGCTATTACGTGTATATACCAATTC 1530
DB 4593 ATGATGGCGCTTTTGTGTCAGATTGTAATTTCAATTAATAGATCAAAATATTTATGCTAATA 4652
QY 1531 GTTTACAGATTCTTCTCTGAATCGTTTGGAGTGGAATCGAGTTTATACAAATAGAT 1590
DB 4653 TTTTCTTATTTCAMACCAAAATAAATGAATTTATTTGCTCTGTGTAATAAAAAA 4712
QY 1591 ACTCAGTTTGTGAATGGTTTAAAAA 1629
DB 4713 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4751

RESULT 5
US-09-925-065A-685225/c
Sequence 685225, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
PRIORITY FILING DATE: 2001-08-08
PRIORITY FILING DATE: 2000-10-24
PRIORITY FILING DATE: 2000-11-20
PRIORITY FILING DATE: 2000-11-20
PRIORITY FILING DATE: 2000-11-30
PRIORITY FILING DATE: 2000-11-30
PRIORITY FILING DATE: 2001-01-16
PRIORITY FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 827921
LENGTH: 615
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-827921

Query Match      2.4%; Score 38.8; DB 6; Length 615;
Best Local Similarity 63.0%; Pred. No. 3;
Matches 58; Conservative 1; Mismatches 33; Indels 0; Gaps 0;

QY 1535 ACAGATTCTCTGTGAATCGTTTGGAGTGGGATCGAGTTTATACAAATAGATCTC 1594
DB 565 AAAAACTCTCTTAAATTTCAATGTTAGAGGAAAGAGTATACAAAGATAATATCTACT 506
QY 1595 AGTTTGAATGGTTTAAAAA 1626
DB 505 TATATTAGTGTATTTTAAAAA 474

RESULT 7
US-09-925-065A-852195
Sequence 852195, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
PRIORITY FILING DATE: 2001-08-08
PRIORITY FILING DATE: 2000-10-24
PRIORITY FILING DATE: 2000-11-20
PRIORITY FILING DATE: 2000-11-20
PRIORITY FILING DATE: 2000-11-30
PRIORITY FILING DATE: 2001-01-16
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 827921
LENGTH: 615
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-827921

Query Match      2.4%; Score 39.4; DB 6; Length 1386;
Best Local Similarity 61.0%; Pred. No. 3;
Matches 64; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 925 AAGGGAAGGAGGAGTGGTGGGAGGATGCCAAAGCTTAAGGAATGTGCTTTCACAAAGG 984
DB 1137 AAGGAGAGGAGAGGAGGAGTGGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGA 1078
QY 985 AAATGTAAAGCAGACAGAGAGGAGGATGTTAATGGAATGGGCT 1029
DB 1077 GAAGGAGGAGAGAGAGAGAGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGA 1033

RESULT 6
US-09-925-065A-827921/c
Sequence 827921, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
PRIORITY FILING DATE: 2001-08-08
PRIORITY FILING DATE: 2000-10-24
PRIORITY FILING DATE: 2000-11-20
PRIORITY FILING DATE: 2000-11-20
PRIORITY FILING DATE: 2000-11-30
PRIORITY FILING DATE: 2000-11-30
PRIORITY FILING DATE: 2001-01-16
PRIORITY FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 827921
LENGTH: 615
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-827921

Query Match      2.4%; Score 38.8; DB 6; Length 615;
Best Local Similarity 63.0%; Pred. No. 3;
Matches 58; Conservative 1; Mismatches 33; Indels 0; Gaps 0;

QY 1535 ACAGATTCTCTGTGAATCGTTTGGAGTGGGATCGAGTTTATACAAATAGATCTC 1594
DB 565 AAAAACTCTCTTAAATTTCAATGTTAGAGGAAAGAGTATACAAAGATAATATCTACT 506
QY 1595 AGTTTGAATGGTTTAAAAA 1626
DB 505 TATATTAGTGTATTTTAAAAA 474

RESULT 7
US-09-925-065A-852195
Sequence 852195, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
PRIORITY FILING DATE: 2001-08-08
PRIORITY FILING DATE: 2000-10-24
PRIORITY FILING DATE: 2000-11-20
PRIORITY FILING DATE: 2000-11-20
PRIORITY FILING DATE: 2000-11-30
PRIORITY FILING DATE: 2001-01-16
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 827921
LENGTH: 615
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-827921
```

```
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925.065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 852195
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-852195

Query Match      2.4%; Score 38.8; DB 6; Length 622;
Best Local Similarity 63.0%; Pred. No. 3;
Matches 58; Conservative 1; Mismatches 33; Indels 0; Gaps 0;

Qy 1535 ACAGATCTCTCTGGAAGTCTGGAAGTGGAGTGGAGTTTATACAAATAAGTACTC 1594
Db 52 AAAAACTCTCTTAAATTTCAATGTTAGAGGAAAGAGTAAACAAGATATATACTC 111

Qy 1595 AGTTTGAATCGTTTAAAAA-AAAAAAAAA 1626
Db 112 TATATTAGTATTTTAA-AAAAAAAAA 143

RESULT 8
US-11-098-686-8738/c
; Sequence 8738, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098.686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8738
; LENGTH: 194553
; TYPE: DNA
; ORGANISM: Lawsonia intracellularis
US-11-098-686-8738

Query Match      2.4%; Score 38.8; DB 12; Length 194553;
Best Local Similarity 48.6%; Pred. No. 48;
Matches 106; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

Qy 1411 ATTGTAAATGCCTTAAGATGGGATATACAGAAATTTCTTCGTTTGGGTGCAATT 1470
Db 70218 ATTTTATATACAAATTATATTTATTTGTTATCTATTTTGTCTATTTT 70159

Qy 1471 CCACATTTGACTATAGTACAGCGGTCTCTTTATGCTATTACGTGTATACCAATC 1530
Db 70158 TGATTATATCTTTAGTTAATACTGATCTAUCITTTTATTTTCTTTATTCATTA 70099

Qy 1531 GTTTACAGATCTTCTGTGAACTCGTTGGAAGTGGAGTGGAGTTTATACAAATAAGAT 1590
Db 70098 GAATTCATATCTTCTTACATAGGGAAGATATGTTCAATATAGAAATTTATCTTAAATATA 70039
```

```
Qy 1591 ACTCAGTTTGAATGGTTTAAAAA-AAAAAAAAA 1628
Db 70038 GAAAAGCTATTATGAAGTTAACTATATACTATATAA 70001

RESULT 9
US-09-925-065A-335502/c
; Sequence 335502, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925.065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 335502
; LENGTH: 569
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-335502

Query Match      2.4%; Score 38.6; DB 6; Length 569;
Best Local Similarity 63.4%; Pred. No. 3.3;
Matches 59; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 1258 TGGAAATTTTAGACGACATGGAGACAGAAAGATTTGCAAGAGTGAACCTGCAATTGGG 1317
Db 224 TTGTAGACTTGGACAAAGATGTTTACAGAAAGGATATTAAGATGTGTACTGGATCTTGT 165

Qy 1318 AAGCTTCCTCCAGATATTTATTTTCCCAAGCA 1350
Db 164 TATCTGTATTTTGATATGTAATGTAATTTCCCAAGGA 132

RESULT 10
US-09-925-065A-687471/c
; Sequence 687471, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925.065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 687471
; LENGTH: 661
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
US-09-925-065A-687471

Query Match      2.3%; Score 38; DB 6; Length 661;
Best Local Similarity 48.6%; Pred. No. 5.1; Indels 0; Gaps 0;
Matches 104; Conservative 0; Mismatches 0;

Qy 1391 GAGTGGATCTCTTGATGGCATTGTAATGCCCTTAAGATGAGGATATCAGGAAATTTTC 1450
Db 230 GTTGTGTTTATGGTTTTCGCAATGTTTAGACATTTAAATATTTATATAGTCAATATTT 171

Qy 1451 TTGCTGTTTGGTTCGCAATCCATTTGACATACGATACAGCGGTTCTTTTATGCT 1510
Db 170 TCATGTTTTCATGTAATGTCCTTCTCAGGATGCTCAGAGAGACCTTCCCAATGTA 111

Qy 1511 ATTACGTTGATATACCAATCTGTTTACAGATTTCTCTGTAAGTCTGTTGGAAGTGGGAATG 1570
Db 110 AGTCGTTTCTCTCTCATTTCTTTAGAGCATCATTTGTAATCTACCTGGAATTCATGTTT 51

Qy 1571 GAGGTTTATACAAATAAGATACTCAGTTTTGAAT 1604
Db 50 GTATTTTGGYAGATTATGCGCTCACAGTATTATT 17
```

```
RESULT 11
US-09-925-065A-159939
; Sequence 159939, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 159939
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-159939
```

```
Query Match      2.3%; Score 37.6; DB 6; Length 535;
Best Local Similarity 57.8%; Pred. No. 5.9;
Matches 67; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 1513 TAGCTGTATACACATTCGTTTACAGATTTCTCTGTAAGTCTGTTGGAAGTGGGAATGGA 1572
Db 368 TATGTTTGTATTTCTTCTCTGTCGTAATGCAATGTACACTTTGGAAATGTTTGTGTC 427

Qy 1573 GGTTTATACAAATAAGATACTCAGTTTTGAATGTTTAAAAAATAAATAAATAA 1628
Db 428 TGGTTTAATAAATTTTATGTTGTCATATTTTAAAGTAATAATAATAACA 483
```

```
RESULT 12
US-09-925-065A-827920/c
; Sequence 827920, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
```

```
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 827920
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-827920
```

```
Query Match      2.3%; Score 37.6; DB 6; Length 615;
Best Local Similarity 63.0%; Pred. No. 6.3;
Matches 58; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 1535 ACAGATTTCTCTGTAAGTCTGTTGGAAGTGGGAATGAGGTTTATACAAATAAGATACTC 1594
Db 565 AAAAATCTCTTTAAATTTCAATGTTAGAGGAAAGAGTAATACAAAGATAATATACTC 506

Qy 1595 AGTTTGAATGTTTAAAAAATAAATAAATAA 1626
Db 505 TATATTTAGTGTATTTTAAAAAATAAATAA 474
```

```
RESULT 13
US-09-925-065A-852196
; Sequence 852196, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 852196
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-852196
```

```
Query Match      2.3%; Score 37.6; DB 6; Length 622;
Best Local Similarity 63.0%; Pred. No. 6.3;
Matches 58; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 1535 ACAGATTTCTCTGTAAGTCTGTTGGAAGTGGGAATGAGGTTTATACAAATAAGATACTC 1594
Db 52 AAAAATCTCTTTAAATTTCAATGTTAGAGGAAAGAGTAATACAAAGATAATATACTC 111

Qy 1595 AGTTTGAATGTTTAAAAAATAAATAAATAA 1626
```

```
Db      112 TATATTAGTGATATTTTAAATAAATAAATAA 143
|||||
Query Match      2.3%; Score 36.8; DB 6; Length 459;
Best Local Similarity 56.7%; Pred. No. 9;
Matches 68; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

RESULT 14
US-09-925-065A-466709/c
; Sequence 466709, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 466709
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-466709

Query Match      2.3%; Score 37; DB 6; Length 519;
Best Local Similarity 67.5%; Pred. No. 8.5;
Matches 52; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy      1553 TCGTGGAGTGGGAGGTTTATACAAATAGATCTCAGTTTGAATGGTTTAA 1612
Db      450 TAGTCTGAGGATGGGTGGAGTCTCTTACAAATAGGAACCCGCTCATCAGCCTTAAA 391

Qy      1613 AAAAAAAAAAAAAAAAAA 1629
Db      390 AAAAAAAAAAAAAAAAAA 374

RESULT 15
US-09-925-065A-192552
; Sequence 192552, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 192552
; LENGTH: 459
; TYPE: DNA
; ORGANISM: Homo sapiens
```

US-09-925-065A-192552

```
Query Match      2.3%; Score 36.8; DB 6; Length 459;
Best Local Similarity 56.7%; Pred. No. 9;
Matches 68; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
```

```
Qy      1510 TATTACGTGTATATACCATTCGTTTACAGATTCTTCTGTGAACCTCGTTGGAAGTGGGAAT 1569
Db      339 TATTCTCTGTGTTTCCGCCCTTTTCCACTGTGTCTTATTTTCCATGCAGCAAAAGTAAAA 398

Qy      1570 GGAGGTTTATACAAATAAGATACCTCAGTTTGAATGGTTTAAAAA 1629
Db      399 AATTTTAAATGAAAAATCAGATTGTGAATGTTTCTCTGTA 458
```

Search completed: February 21, 2006, 02:35:12  
Job time : 541 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 20, 2006, 23:16:16 ; Search time 427 Seconds  
(without alignments)  
2316.700 Million cell updates/sec

Title: US-10-731-525-8

Perfect score: 2445

Sequence: 1 MALAFTSSAPTSTPPSSE.....KYPSEKRDVDSLMAFVNALR 466

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 7204323 seqs, 1061406715 residues

Total number of hits satisfying chosen parameters: 14408646

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DBV=xlh  
-O=/abs/ABSSWEB\_spool/US10731525/runat\_17022006\_164911\_411/app\_query.fasta.1  
-DB=Published Applications\_NA\_New -QWMT=fascap -SUFFIX=p2n.rnpbn -MINMATCH=0.1  
-LOOPCU=0 -LOOPTXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -HOST=abes03h  
-USER=US10731525 @CGN 1.1.220 @runat\_17022006\_164911\_411 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DISPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New:

1: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq1.\*  
7: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq1.\*  
9: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq2.\*  
11: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq3.\*  
12: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq4.\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1745.5	71.4	1082144	12	US-11-117-187-211 Sequence 211, App
2	307.5	12.6	4055	8	US-10-793-626-3357 Sequence 3357, App
3	250.5	10.2	468	8	US-10-793-626-307 Sequence 307, App
4	170	7.0	299	12	US-11-128-061-705 Sequence 705, App

5	170	7.0	299	12	US-11-128-061-4347	Sequence 4347, App
6	170	7.0	299	12	US-11-128-049-705	Sequence 705, App
7	170	7.0	299	12	US-11-128-049-4347	Sequence 4347, App
8	159.5	6.5	1938	8	US-10-821-234-557	Sequence 557, App
9	139.5	5.7	4040	8	US-10-131-826A-425	Sequence 425, App
10	131	5.4	3155	8	US-10-793-626-3424	Sequence 3424, App
11	129.5	5.3	2578	12	US-11-136-527-219	Sequence 219, App
12	127.5	5.2	1518	8	US-10-821-234-435	Sequence 435, App
13	126	5.2	2514	12	US-11-122-329-23	Sequence 23, Appl
14	123.5	5.1	5801	12	US-11-000-463-580	Sequence 580, App
15	118	4.8	1323	8	US-10-821-234-430	Sequence 430, App
16	118	4.8	2344	8	US-10-947-249-126	Sequence 126, App
17	116	4.7	2206	12	US-11-128-061-554	Sequence 554, App
18	116	4.7	2206	12	US-11-128-049-554	Sequence 554, App
19	110.5	4.5	3133	12	US-11-000-463-108	Sequence 108, App
20	109	4.5	1881	8	US-10-467-657-5431	Sequence 5431, App
21	107	4.4	7231	12	US-11-136-527-2622	Sequence 2622, App
22	106.5	4.4	1400	12	US-11-136-527-6314	Sequence 6314, App
23	106.5	4.4	1896	12	US-11-136-527-2418	Sequence 2418, App
24	104	4.3	2712	8	US-10-821-234-8	Sequence 8, Appli
25	99	4.0	4119	12	US-11-128-061-1127	Sequence 1127, App
26	99	4.0	4119	12	US-11-128-049-1127	Sequence 1127, App
27	99	4.0	4170	12	US-11-024-959-92	Sequence 92, Appl
28	94	3.8	991	8	US-10-750-185-30544	Sequence 30544, A
29	94	3.8	991	8	US-10-750-623-30544	Sequence 30544, A
30	92.5	3.8	7659	12	US-11-091-883-498	Sequence 498, App
31	92.5	3.8	11151	12	US-11-052-554A-525	Sequence 525, App
32	92	3.8	1344	8	US-10-821-234-732	Sequence 732, App
33	92	3.8	2131	8	US-10-775-169-246	Sequence 246, App
34	91.5	3.7	552	8	US-10-467-962B-5	Sequence 5, Appli
35	91.5	3.7	4874	8	US-10-995-561-94	Sequence 94, Appl
36	90.5	3.7	2346	6	US-09-925-065A-1321	Sequence 1321, App
37	90.5	3.7	4089	8	US-10-775-169-137	Sequence 137, App
38	90.5	3.7	5053	8	US-10-995-561-93	Sequence 93, Appl
39	90.5	3.7	137671	12	US-11-121-086-47	Sequence 47, Appl
40	90	3.7	495	6	US-09-925-065A-66769	Sequence 66769, A
41	89.5	3.7	2277	12	US-11-059-275-1	Sequence 1, Appli
42	89.5	3.7	2283	8	US-10-858-730-147	Sequence 147, App
43	89.5	3.7	5676	12	US-11-134-795-5	Sequence 5, Appli
44	89.5	3.7	6221	12	US-11-134-795-6	Sequence 6, Appli
45	89	3.6	5534	12	US-11-040-472-11	Sequence 11, Appl

#### ALIGNMENTS

RESULT 1

US-11-117-187-211  
; Sequence 211, Application US/11117187  
; Publication No. US20050266560A1  
; GENERAL INFORMATION:  
; APPLICANT: PREUSS, DAPHNE  
; APPLICANT: COPENHAVER, GREGORY  
; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS  
; FILE REFERENCE: ARCD:309US  
; CURRENT APPLICATION NUMBER: US/11/117,187  
; CURRENT FILING DATE: 2005-04-28  
; PRIOR APPLICATION NUMBER: US/09/531,120  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: 60/125,219  
; PRIOR FILING DATE: 1999-03-18  
; NUMBER OF SEQ IDS: 212  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 211  
; LENGTH: 1082144  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; US-11-117-187-211

Alignment Scores:  
Pred. No.: 4.12e-154 Length: 1082144  
Score: 1745.5 Matches: 342  
Percent Similarity: 80.5% Conservative: 43  
Best Local Similarity: 71.5% Mismatch: 47









```
US-11-128-049-4347
; Sequence 4347, Application US/11128049
; Publication No. US20060010513A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounse, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FOR
; FILE OF INVENTION: MAKING AND USING SAME
; FILE REFERENCES: 01997.027700
; CURRENT APPLICATION NUMBER: US/11/128,049
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4347
; LENGTH: 299
; TYPE: DNA
; ORGANISM: Cricetulus griseus
; US-11-128-049-4347

Alignment Scores:
Pred. No.: 3,82e-07 Length: 299
Score: 170.00 Matches: 39
Percent Similarity: 56.0% Conservative: 22
Best Local Similarity: 35.8% Mismatches: 37
Query Match: 7.0% Indels: 11
DB: 12 Gaps: 3

US-10-731-525-8 (1-466) x US-11-128-049-4347 (1-299)
QY 198 LysValArgProLeuArgAlaLeuLysGlyLeuArgAlaTrpIleThrGlyGlnArg 217
Db 4 AACTAGAACCTTTAAGACGCTCTTTAAACGGTAATCAGCTATGGGTACCGGTATCAGA 63

QY 218 LysAspGlnSerProGlyThrArgSerGluIleProValValGlnValAspProAlaPhe 237
Db 64 TCTGACCAATCTGTA---AACAGCAGCATATGACCAACCTGGAGTGGATGAGCAA--- 117

QY 238 GluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsnProValAlaAsnValLys 257
Db 118 -----AATCAGTTGATCAAGTTTCATCCGATCTTTTCTGGTCG 156

QY 258 GlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAsnSerLeuHisAla 277
Db 157 CTTGATGAGGTGAAGGAATACATTAAAGAAATATATTGTTTACACACATTACATGAT 216

QY 278 LysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeuProGlyGlnHis 297
Db 217 AAGGGTTTCCTAGTAGTAGTGTCTCTTGTGTAAGACGACGTAACACAGGGCGAG--- 273

QY 298 GluArgGlu-GlyArgTrpTrp 305
Db 274 GATTTAGAGCTGCGGTGGTTGTGG 298

RESULT 8
US-10-821-234-557
; Sequence 557, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Grain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
```

```
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: PC_SEQ_genes Version 1.0
; SEQ ID NO 557
; LENGTH: 1938
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-821-234-557

Alignment Scores:
Pred. No.: 3,17e-05 Length: 1938
Score: 159.50 Matches: 112
Percent Similarity: 33.7% Conservative: 81
Best Local Similarity: 19.6% Mismatches: 206
Query Match: 6.5% Indels: 174
DB: 8 Gaps: 23

US-10-731-525-8 (1-466) x US-10-821-234-557 (1-1938)
QY 7 SerSerIleSerAlaProThrSerThrPheProSerSerGluProLysLeuProGlnIle 26
Db 352 GCCAAGATCGATGCAACCTCAGCGTCTGTGTCGCCAGCAGAGTTTGATGTGATGGGTAC 411

QY 27 GlySerIleArgIleSerGluArgProIleGlyGlyAlaValAsnPheAsnLeuSerGln 46
Db 412 CCCACCATCAAGATCCTTAAGAAG-----GGCAGGCTGTAGACTACGAGGGCTCCAGA 465

QY 47 ArgArgSerLeuValLysProValAsnAlaGluProProArgLysAspSerIleValPro 66
Db 466 ACCAGGAAGAAATTTGTCACAGGTGTCAGAGAGTCTCCAGCGCCGACTGGAGCGCTCCA 525

QY 67 LeuAlaAlaThrIleValAlaSerAlaSerGluThrLysGluGluAspPheGluGln 86
Db 526 CCAAGAGTCACGCTGTGTGTGACC-----AAAGAGAACTTTGATGAA 567

QY 87 IleAlaSerAspLeuAsp----- 92
Db 568 GTTGTGATGATGCAGATATCATCTCTGTGGAGTTTTTATGCCCATGGTGTGGACACTGC 627

QY 93 -----AsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAsp 110
Db 628 AAGAAACTTGCCTCCCGCAGTATGAGAGAGCGCCAGAGGAGCTCAGAACGCTTCTCTCTCCA 687

QY 111 IleAlaIleAla---PheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeu 129
Db 688 ATTCCCTGGCAAGGTTCAGCCGCCAGAAACAGACCTGGCCCAAGAGGTTTGTATGTC 747

QY 130 ThrGlyArgPro---PheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThr 148
Db 748 TCTGGGTATCCACCTCGAAATTTTCCGC-----AAGGAAGG-----CCTTATGAC 795

QY 149 TyrGlnLeuPheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPhe--- 167
Db 796 TAC-----AACGGCCACGAGAAAAATATGGAATC---GTTGATTACATGATCGAG 843

QY 168 -----ProAspAlaValGluValGlnAlaLeuValArgSerLysGlyLeuPheSer 184
Db 844 CAGTCCGGGCTCTCCCTCCAGAGGAGATTCTGACCCTG-----AAGCAGGTCAGGAG 894

QY 185 PheTyrGluAspGlyHisGlnGluCysCysArgValArgLysValArgProLeuArgArg 204
Db 895 TTCCTGAAGGATGGAGACGAT----- 915

QY 205 AlaLeuLysGlyLeuArgAlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThr 224
Db 916 -----GTCATCATCATCGGGGTCTTTAAGGGGAGAGT----- 948

QY 225 ArgSerGluIleProValGlnValAspProAlaPheGluGlyMetAspGlyGlyIle 244
Db 949 -----GACCCAGCTTACCAACATACCAAGGATGCCGCT 981
```

```
Qy 245 GlySerLeuValLysTrpAsnProValAlaAsnValLysGlyHisPheTyrAsnPhe 264
Db 982 AACAACTGAGAGAGATTACAATTTCCACACACTTTCAGCACAGAAATAGCAAAATTC 1041
Qy 265 LeuArgThrMetAsnValProValAsnSerLeuHisAlaLysGlyTyrValSer----- 282
Db 1042 TTGAAGTCTCCAGGGCAGTTGGTTGTAATGAGCGCTCGAGAAATCCAGTCCAGTAT 1101
Qy 283 -----Ile-GlyCysGluProCysThrArgPro----- 291
Db 1102 GAGCCCGGAGCCACATGATGAGCGTCCAGGGTCCACCCAGGACTCGGCCATCAAGGAC 1161
Qy 291 ----- 291
Db 1162 TTGCTGTGAAGTACGCCCTGCTGGTGGCCACCGCAAGGTGTCAACGATGCTAAG 1221
Qy 292 -ValLeuProGlyClnHisGluArgGluGlyArgTyrTip-----TrpG1 306
Db 1222 CGCTACACAGGC-----GCCCTGTGTGTCTACTACTACGTGTGGAC 1266
Qy 306 uAspAlaLysAlaLysGluCysGlyLeuHisLysGlyAsnValLysGln----- 322
Db 1267 TTCAGCTTTGATTACAGAGCTGCAACTCAGTTTGGCGGAGCAAAAGTCTAGAGGTGCC 1326
Qy 323 -----GlnLysGluGluAspValAsnGlyAsn-- 331
Db 1327 AAGACTTCCCTGAGTACACTTTGCCATTGCGGAC--GAGAGGACTATCTGGGGAGGT 1385
Qy 332 -----GlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrVal----- 345
Db 1386 GAAGGACTGGGCTCAGCAGAGTGGGGAGGATGCTAATGCCCATCTCGGACGAGAG 1445
Qy 346 -----ProAspIlePheAsnSerProAsn-----Va 354
Db 1446 TGGGAAGAAGTTGCCATGAGCAGAGGATTTGACTCTGACACCCCTCGCGAGTTGT 1505
Qy 354 lValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGlu----- 369
Db 1506 CACTGCTTTCAAAAAGGAAACTGAAGCCAGTCAATCAATCCAGCCAGTGCCTCAAGAA 1565
Qy 370 -----AspAr 371
Db 1566 CAACAAGGACCGCTCAAGTCTGTGGTGGGAAGACCTTTGACTCTCCATTGTGAGACC 1625
Qy 371 gLysGluProTrpLeuValLeuTyrAlaProTyrCysProTyrCysGlnAlaMetG1 391
Db 1626 CAAGAAGGAGCTCTCATCGAGTTCTACGACCATGTTGGTGGGCGACTGCAAGCAGCTAGA 1685
Qy 391 uGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLy 411
Db 1686 GCCCGTGTACAACAGCTGTGGCCAAAGAGTACAAGGGCCAAAGGGCTGTGTATCGCCAA 1745
Qy 411 sPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLysSerPh 431
Db 1746 GATG-----GACGCCACTGCCACGACGCTCCCGAGCGACCGCTATAAGTGGAGGCTT 1799
Qy 431 eProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSerG1 451
Db 1800 CCCCACCATCTACTCTCGCCCCCAGTGGGGGACAAAAGAACCCAGTTAAATTTGAGGGTGG 1859
Qy 451 uLysArgAspValAspSerLeuMetAlaPheVal 462
Db 1860 AGACAGAGATCTGGAGCATTTTGACAAAGTTTATA 1893
```

## RESULT 9

```
US-10-131-826A-425
; Sequence 425, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Deanoyers, Luc
```

```
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: F3330R1C128
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 425
; LENGTH: 4040
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-10-131-826A-425
```

```
Alignment Scores:
Pred. No.: 0.00599 Length: 4040
Score: 139.50 Matches: 113
Percent Similarity: 37.6% Conservative: 68
Best Local Similarity: 23.5% Mismatches: 209
Query Match: 5.7% Indels: 92
DB: 8 Gaps: 24
```

US-10-731-525-8 (1-466) x US-10-131-826A-425 (1-4040)

```
Qy 27 GlySerIleArgIleSerGluArgProIleGlyAlaValAsnPheAsnLeuSerCln 46
Db 932 GGGTTACTTCGAAT-----GGAGCTGTAACTGTGTGTATGATAGA 973
Qy 47 ArgArgSerLeuValLysProValAsnAlaGluProArgLysAspSerIleValPro 66
Db 974 ATGCTTTGCCGAATGAAGAGGATCAACAGTATCCC----- 1009
Qy 67 LeuAlaAlaThrThrIleValAlaSerAlaSerGluThrLysGluIuAspPheGluCln 86
Db 1010 -----AGTCTCTTCATTTTTCGGTCTGG-AATGGCCCCAGTGAATATCATGGAGACAG 1062
Qy 87 IleAla-----SerAspLeuAspAsnAlaSerProLeuGluIleMetAspArgAla 103
Db 1063 ATCAAGGAGAGGTTTACTGAGTTTTCACATTCAGACTGTTAGTAGTACAGTACAGCAACT 1122
Qy 104 LeuAsp---Lys-PheGlyAsnAspIleAlaIleAlaPheSer----- 116
```



ORGANISM: Rattus norvegicus  
US-11-136-527-219

## Alignment Scores:

Pred. No.: 0.0328 Length: 2578  
Score: 129.50 Matches: 114  
Percent Similarity: 33.5% Conservatives: 77  
Best Local Similarity: 20.0% Mismatches: 211  
Query Match: 5.3% Indels: 170  
DB: 12 Gaps: 23

US-10-731-525-8 (1-466) x US-11-136-527-219 (1-2578)

```
Qy 7 SerSerileSerAlaProThrSerPheProSerSerGluProLysLeuProGlnIle 26
Db 634 GCAAGATGACGCGACTCAGCATCCATGTTGGCCAGCAATTCGATGTGAGTGGGTAC 693
Qy 27 GlySerileArgileSerGluArgProIleGlyAlaValAanPheAan----- 43
Db 694 CCCACCATCAAGATCCTGAAGAG-----GGACAGGCTGTGACTATGACGGCTCCAGG 747
Qy 44 -----LeuSerGlnArgArgSerLeuValLysProValAanAlaGluPro 58
Db 748 ACCCAGGAAGAAATGTTGCCAAAGTCAGAGAACTTCCAGGCTGATGGACACCTCCA 807
Qy 59 ProArgLys-----AspSerileValProLeuAlaAla 69
Db 808 CCGAAGTCACTTACATTGACTAAAGAGAACTTTCAGCATGTGTAAATAATGACAGAC 867
Qy 70 ThrThrIleVal-----AlaSerAlaSerGluThrLysGluGluAsp 83
Db 868 ATCATTGTTGGAGTCTATGCCCCATGTCGGGACACTGCAAGAACTTGGCCCTGAG 927
Qy 84 PheGluGlnIleAlaSerAspLeuAspAanAlaSerProLeuGluIleMetAspArgAla 103
Db 928 TATCAGAGGCTGCCAAGSWSAGCAAGCGCTCTCCCCCAATTCCTTAGCCAAAGATT 987
Qy 104 LeuAspLysPheGlyAanAspIleAlaPheSerGlyAlaGluAspValAlaLeu 123
Db 988 GATCCACCGNACAGACAGACCTCGCTAAGAGTTT-----GATGCTCTGGC 1035
Qy 124 IleGluTyrAlaLysLeu-----ThrGlyArgProPheArgValPheSerLeuAspThr 141
Db 1036 TACCCCACTTTGAAATATTCGCAAGAGGAGGCTTTTGAC----- 1077
Qy 142 GlyArgLeuAanProGluThrTyrGlnLeuPheAspAlaValGluLysHisTyrGlyIle 161
Db 1078 -----TACAATGGTCCACGAGAGAAATATGGAATT 1107
Qy 162 ArgIleGluTyrMetPhe-----ProAspAlaValGluValGlnAlaLeu-Va 177
Db 1108 ---GTTGACTACATGGTTGACAGCTCTGGACCTCTCTCCAAAGGAGATTCTGACCCCTGAAG 1164
Qy 177 lArgSerLys-----GlyLeuPheSerPheTyrGl 187
Db 1165 CAGTCCAGGAGTTCTGAAGGATGGAGATGATGATGATCTCTTGGAGTCTTTGAGGA 1224
Qy 187 uAspGlyHisGlnGluCysCysArgValArgLys---ValArgProLeuArg----- 204
Db 1225 GTCGCGWACCAGCGCTACTTCAGTATCAGGATGCTGCTAAACACCTGAGAGAAGATTAC 1284
Qy 205 -----AlaLeuLysGlyLeuArgAla----- 211
Db 1285 AAATTTACCACACTTTCAGACACTGAAATAGCTAGTTCTCGAAAGTCTCCCTGGGAAA 1344
Qy 212 -TriPleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValVa 231
Db 1345 CTGGTGCTGACGACCTGARAAGTTCCAGTCCAAGTAT-GAGCCAGGATGATGTGAT 1403
Qy 231 lGlnValAspProAlaPheGluGluMetAspGlyGlyIleGlySerLeuValLysTrpAs 251
Db 1404 GGATGTTTCAGGGCTCCACAGAG----- 1425
```

```
Qy 251 nProValAlaAanValLysGlyHisAspIleTrpAanPheLeuArgThrMetAanValPr 271
Db 1426 ----GCATCAGCCATTAAAGCATATGTGTGTAACATGACTTGCCTAGTGGGCCACAG 1481
Qy 271 oValAanSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgPr 291
Db 1482 GAAGACTTCTTAACGATGCTAAGCGGTAC-----AGCAAGCGCCC 1520
Qy 291 oValLeuProGlyGlnHisGlu-----ArgGluGlyArgTyrTr 304
Db 1521 CTTGGTGGTTGTATATACAGCGTGACTTCAGCTTTGATTACAGAACTGCTACTCAGTT 1580
Qy 304 pTTP----- 305
Db 1581 TTGGCGTAACAAAGTCTCAGAGGTGGCCAAAGGACTTCCCTGAGTACACATTGCGCAATTGC 1640
Qy 306 -----GluAsp-----AlaLysAlaLysGluCysGlyLeuHisLysGlyAanVally 321
Db 1641 TGATGAAGAAGACTATGCCACAGAGGTGAAGGACCTGGGACTG----- 1683
Qy 321 sGlnGlnLysGluGluAspValAanGlyAanGlyLeuSerGlnSer----- 336
Db 1684 -AGTGAGAGTGGAGAGAGACTCAATGACGCCATCTTAGATGAGAGTGGGAAGAAGTTTGC 1742
Qy 337 -----HisAlaAanGlyAspAla----- 342
Db 1743 CATGGAGCCAGAGGAGTTTGCATTCAGATGCTCTCCGAGAATTTGTCTGCTTCAAAA 1802
Qy 343 -----ThrThrValProAspIlePheAanSerProAs 353
Db 1803 AGGAAACTAAAGCCAGTTATCAATCACAGCCAGTTTCCCAAGAACCAACAGGACCAGT 1862
Qy 353 nValValAanLeuSerArgThrGlyIleGluAanLeuAlaLysLeuGluAspArgLysGl 373
Db 1863 CAGGTTGGTAGTGGGGAAGACCTTTGAT-----GCCATTGTGATGGACCCCAAAA 1913
Qy 373 uProTrpLeuValLeuTyrAlaProTrpCysProTyrCysGlnAlaMetGluGluSe 393
Db 1914 GGAGCTCTCATTAATTCATGCACCGTGTGTGGGCACCTGMRRCGACGTGGAGCCTGT 1973
Qy 393 rTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLysPheAr 413
Db 1974 CTACACAGCCTAGGCAAGAAGTACAAAGGGCCAGAGGACTTGTCTATCGCAAGATG-- 2031
Qy 413 gAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPheProTh 433
Db 2032 ----GATGCTACTCCCAATGACATCACCAAGCCGATACAGSWGAGGGCTTCCCAC 2087
Qy 433 rIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSerGluLysAr 453
Db 2088 CATCTATTTTGGCCCGAGTGGGACAAAAGAACCCCAATTAAGTTTGAGGGTGGCAACAG 2147
Qy 453 gAspValAspSerLeuMetAlaPheValAan 463
Db 2148 AGATYTCGAGCATTTGAGCAAGTTTATCGAT 2178
```

## RESULT 12

US-10-821-234-435  
; Sequence 435, Application US/10821234  
; Publication No. US20050255114A1  
; GENERAL INFORMATION:  
; APPLICANT: Labat, Ivan  
; APPLICANT: Stache-Crain, Birgit  
; APPLICANT: Andarmani, Susan  
; APPLICANT: Tang, Y. Tom  
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia  
; FILE REFERENCE: 821A  
; CURRENT APPLICATION NUMBER: US/10/821,234  
; CURRENT FILING DATE: 2004-04-07  
; PRIOR APPLICATION NUMBER: US 60/462,047  
; PRIOR FILING DATE: 2003-04-07  
; NUMBER OF SEQ ID NOS: 1704  
; SOFTWARE: pt\_SEQ\_genes Version 1.0



10







GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2006, 14:30:00 ; Search time 187 Seconds  
(without alignments)  
1094.923 Million cell updates/sec

Title: US-10-731-525-8

Perfect score: 2445

Sequence: 1 MALAFTSSISAPSTTPFSSE.....KYPSEKRDVDSLAFVNALR 466

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_21.\*

- 1: Geneseq1980s.\*
- 2: Geneseq1990s.\*
- 3: Geneseq2000s.\*
- 4: Geneseq2001s.\*
- 5: Geneseq2002s.\*
- 6: Geneseq2003as.\*
- 7: Geneseq2003bs.\*
- 8: Geneseq2004s.\*
- 9: Geneseq2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2445	100.0	466	3	AAY44745 Soybean A
2	1951	79.8	475	8	ADK74624 Plant ful
3	1923	78.7	472	3	AAY44744 Soybean A
4	1811.5	74.1	465	3	AAG42684 Arabidops
5	1811.5	74.1	465	5	ABB92842 Herbicida
6	1811.5	74.1	465	8	ADN74399 Thale cre
7	1811.5	74.1	487	3	AAG42683 Arabidops
8	1810	74.0	463	3	AAG42685 Arabidops
9	1806.5	73.9	465	3	AAG17338 Arabidops
10	1806.5	73.9	487	3	AAG17337 Arabidops
11	1805	73.8	463	3	AAG17339 Arabidops
12	1801	73.7	458	3	AAG20363 Arabidops
13	1801	73.7	483	3	AAG20362 Arabidops
14	1797	73.5	458	3	AAG32469 Arabidops
15	1758.5	71.9	455	3	AAG04414 Arabidops
16	1738	71.1	454	3	AAG45534 Arabidops
17	1738	71.1	454	5	ABB91430 Arabidops
18	1738	71.1	454	8	ADN73975 Thale cre
19	1687.5	69.0	407	3	AAG04415 Arabidops
20	1670	68.3	406	3	AAG45535 Arabidops
21	1645.5	67.3	463	3	AAY44746 Wheat ade
22	1628.5	66.6	358	3	AAG20364 Arabidops
23	1624.5	66.4	358	3	AAG32470 Arabidops
24	1622.5	66.4	473	8	ADT58251 Plant pol

#### ALIGNMENTS

RESULT 1

AAY44745

ID AAY44745 standard; protein; 466 AA.

XX AAY44745;

DT 04-MAY-2000 (first entry)

DE Soybean APS reductase encoded by cDNA clone sl2.pk0064.g4.

XX Adenosine-5'-phosphosulfate reductase; APS reductase; soybean;

KW sulphate assimilation protein; SAP; clone sl2.pk0064.g4.

XX Glycine max.

XX WO200004161-A1.

XX 27-JAN-2000.

XX 13-JUL-1999; 99WO-US015808.

XX 14-JUL-1998; 98US-0092833P.

XX (DUPO ) DU PONT DE NEMOURS & CO E I.

XX Falco SC, Allen SM;

XX WPI; 2000-182427/16.

XX N-PSDB; AA250066.

PT Nucleic acid fragments encoding sulfate assimilation proteins in plants and seeds useful as probes for isolating cDNAs and genes encoding homologous proteins, and in producing transgenic plants.

PT Claim 1; Fig 1; 47pp; English.

XX The present sequence is soybean adenosine-5'-phosphosulfate (APS) reductase, a sulphate assimilation protein (SAP). This is encoded by cDNA clone designated sl2.pk0064.g4. APS reductase polypeptides are useful for producing antibodies for screening cDNA expression libraries to isolate full length APS reductase cDNA clones. Chimeric genes comprising the APS reductase gene when transformed in a host cell can alter the level of SAP expression. This can be used to increase sulphur containing compounds in the cell, including the nutritionally critical amino acids cysteine and methionine

XX Sequence 466 AA;

SQ



AA44744 standard; protein; 472 AA.  
AA44744;  
04-MAY-2000 (first entry)  
Soybean APS reductase encoded by cDNA clone se4.1lg09.  
Adenosine-5'-phosphosulfate reductase; APS reductase; soybean;  
sulphate assimilation protein; SAP; clone se4.1lg09.  
Glycine max.  
WO200004161-A1.  
27-JAN-2000.  
13-JUL-1999; 99WO-US015808.  
14-JUL-1998; 98US-0092833P.  
(DUPO ) DU PONT DE NEMOURS & CO E I.  
Falco SC, Allen SM;  
WPI; 2000-182427/16.  
N-PSDB; AAZ50065.  
Nucleic acid fragments encoding sulfate assimilation proteins in plants  
and seeds useful as probes for isolating cDNAs and genes encoding  
PT homologous proteins, and in producing transgenic plants.  
Claim 1; Fig 1; 47pp; English.  
The present sequence is soybean adenosine-5'-phosphosulfate (APS)  
reductase, a sulphate assimilation protein (SAP). This is encoded by cDNA  
clone designated se4.1lg09. APS reductase polypeptides are useful for  
CC producing antibodies for screening cDNA expression libraries to isolate  
CC full length APS reductase cDNA clones. Chimeric genes comprising the APS  
CC reductase gene when transformed in a host cell can alter the level of SAP  
CC expression. This can be used to increase sulphur containing compounds in  
CC the cell, including the nutritionally critical amino acids cysteine and  
CC methionine  
XX  
SQ Sequence 472 AA;  
Query Match 78.7%; Score 1923; DB 3; Length 472;  
Best Local Similarity 79.4%; Pred. No. 3.5e-189;  
Matches 378; Conservative 37; Mismatches 45; Indels 16; Gaps 7;  
QY 1 MALAFTSSIS-----APTSTF-----PSSBPKLPQIGSIRISRPPIGGAVNENLSORRSL 50  
DB 3 LAVSTSSSSAAAAAASSFFRLGSSSDAKAPQIGSFRPFPQVSSGVNLTORRSL 62  
QY 51 VKPVNAEPPRKOSIVPLAATTIVASASETKBEDFEQIASDLDNASPLEIMDRALDKFGND 110  
DB 63 VRPLNAEPQRNSDVPVPLAA-TI VAPVEKEKEDEFEQLAKDLENSPLEIMDKALERFGND 121  
QY 111 IAIAFGAEDVALIEYAKLTGRPRVPSLDTGLNPTETVQLFDPAVEKHGIRIENYFPDA 170  
DB 122 IAIAFGAEDVALIEYAHLTGRPRVPSLDTGLNPTETVYKFFDAVEKHGIRIENYFPDA 181  
QY 171 VEQVALVRSGKLSFFEDGHOECRVRKVRPLRLALKGLRAWITGORKQDSPTGSEIPV 230  
DB 182 VEQVALVRKGLSPFEDGHOECRVRKVRPLRLALKGLKAWITGORKQDSPTGSEIFI 241  
QY 231 VQVDAPEGMGGIGSLVKWNPVANKGHDIWNFLRTMNPVNSLHAKGVSVSGCEPCTR 290  
DB 242 VQVDPVFEGLDGGIGSLVKWNPVANKGLDIWNFLRTMNPVNSLHAKGVSVSGCEPCTR 301  
QY 291 PVLPGQHERGRRWWEDAKAKCEGLHKNVKKQKEEDVNGSLQSHANGDATTVPDIEN 350  
DB 302 PVLPGQHERGRRWWEDAKAKCEGLHKNLKBQDAALQNGNGTSQ--GNGSA-TVADIPI 358

QY 351 SPNVNLSRTGIENLAKLEDRKEPWLVLVYAPWCPYCOAMEESYVDLADKLAGSTGMKV 410  
DB 359 SQNVVSLRSGIENLAKLENRKEHVLVLVYAPWCRFCQAMEESYVDLAELKARS-GYKVA 417  
QY 411 KFRADGEQKEFAKSELQGLSFPFLLFPKHSSRPTIKYPSEKRDVDSLMAFVNALR 466  
DB 418 KFRADGEQKEYAKSELQGLSFPFLLFPKHSSQP-IKYPSEKRDVDSLTAFFVNALR 472

RESULT 4  
AAG42684  
ID AAG42684 standard; protein; 465 AA.  
XX  
AC AAG42684;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 53260.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 29-MAR-1999; 99US-0126264P.  
PR 01-APR-1999; 99US-0126785P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.





Db 1 MAMSVNVSSSSSGIINSRFGVSLPKVQIGSLRLLLDRVHVAPVSNLSSGKSSSSVKPL 60  
 QY 55 NABPPRRKDSIVPLAAATTVIVASSETKE---EDFEQIASDLDNASPLRIMDRALDKFGNDI 111  
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 120  
 QY 112 AIAFSGAEDVALIEYAKLTGRPPRVFSLDTGRNLNPETYQLFDAVEKHGIRIERYMFPDVA 171  
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 180  
 QY 172 EVQALVRSKGLFSFYEDGHQECRCRVKVRPLRRALKGLRAWITGORKQDSPTGTRSEIPV 231  
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 240  
 QY 232 QVDPAPFEGMDGGISLVKWNPNVANVKGHDINWFLRTMNVVPSNLSHAKGYVSIIGCEPCTRP 291  
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 300  
 QY 292 VLPQHERBGRWWEDAKACEGLHKGNVKQKEEDVNGLSQSHANGDATTVPDIFNS 351  
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 352  
 QY 352 PNVVNLRTGIEINLAKLEDRKEPWLVLVYAPWCPYCOAMEESYVDLADKLKAGSTGMKVGK 411  
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 411  
 QY 412 FRADGQKEFAKQELQSGSPFTILFFPKHSRPTIKYPSEKRDVDSLMAFVNLAR 466  
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 465

## RESULT 5

ID ABB92842 standard; protein; 465 AA.

AC ABB92842;

XX 31-MAY-2002 (first entry)

XX Herbicidally active polypeptide SEQ ID NO 2053.

XX Herbicidal; plant; agriculture; herbicide.

XX Arabidopsis thaliana.

XX WO200210210-A2.

XX 07-FEB-2002.

XX 28-AUG-2001; 2001WO-EP009892.

XX 28-AUG-2001; 2001WO-EP009892.

XX (FARB ) BAYER AG.

XX Tietjen K, Weidler M;

XX WPI; 2002-269010/31.

XX Identifying plant target proteins for herbicidally active compounds,  
 PT comprising aligning and comparing nucleic acid or amino acid sequences  
 PT from plant with nucleic acid or amino acid sequences from non-plant  
 PT organisms.

XX Claim 5; SEQ ID NO 2053; 261pp + Sequence Listing; English.

XX The invention relates to identifying target proteins (ABB90790-ABB94016)  
 CC for herbicidally active compounds, comprising aligning and comparing  
 CC nucleic acid or amino acid sequences from plant with nucleic acid or  
 CC amino acid sequences from non-plant organisms using suitable search  
 CC parameters, where plant sequences having an E-value greater by a factor  
 CC of 3 than the E-value of most similar non-plant sequences are selected.  
 CC The polypeptides or nucleic acids encoding them are useful for

CC identifying modulators. The identified modulators are useful as  
 CC herbicides

SQ Sequence 465 AA;

Query Match 74.1%; Score 1811.5; DB 5; Length 465;  
 Best Local Similarity 73.9%; Pred. No. 1.2e-177;  
 Matches 351; Conservative 48; Mismatches 57; Indels 19; Gaps 7;

QY 1 MALAFSTTSISAPTSTPSS-----EPKLPOIGSIRISERPIGAGVNFNLS-QRRSLVKPV 54  
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 60

QY 55 NABPPRRKDSIVPLAAATTVIVASSETKE---EDFEQIASDLDNASPLRIMDRALDKFGNDI 111  
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 120

QY 112 AIAFSGAEDVALIEYAKLTGRPPRVFSLDTGRNLNPETYQLFDAVEKHGIRIERYMFPDVA 171  
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 180

QY 172 EVQALVRSKGLFSFYEDGHQECRCRVKVRPLRRALKGLRAWITGORKQDSPTGTRSEIPV 231  
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 240

QY 232 QVDPAPFEGMDGGISLVKWNPNVANVKGHDINWFLRTMNVVPSNLSHAKGYVSIIGCEPCTRP 291  
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 300

QY 292 VLPQHERBGRWWEDAKACEGLHKGNVKQKEEDVNGLSQSHANGDATTVPDIFNS 351  
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 352

QY 352 PNVVNLRTGIEINLAKLEDRKEPWLVLVYAPWCPYCOAMEESYVDLADKLKAGSTGMKVGK 411  
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 411

QY 412 FRADGQKEFAKQELQSGSPFTILFFPKHSRPTIKYPSEKRDVDSLMAFVNLAR 466  
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 465

## RESULT 6

ADN74399

ID ADN74399 standard; protein; 465 AA.

AC ADN74399;

XX 15-JUL-2004 (first entry)

XX Thale cress protein repressed in E2Fa/Dpa expressing plants SeqID 2294.

XX plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;  
 KW animal feed product; thale cress; cell wall biosynthesis;  
 KW nitrogen metabolism; carbon metabolism.

XX Arabidopsis thaliana.

XX WO2004035798-A2.

XX 29-APR-2004.

XX 20-OCT-2003; 2003WO-EP011658.

XX 18-OCT-2002; 2002EP-00079408.

XX (CROP-) CROPDESIGN NV.

XX Inze D, De Veylder L, Vlieghe K;

XX WPI; 2004-348466/32.

XX N-PSDB; ADN74399.

XX



PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151080P.
PR	22-JUN-1999;	99US-0139899P.	PR	30-AUG-1999;	99US-0151303P.
PR	23-JUN-1999;	99US-0140353P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-0140695P.	PR	07-SEP-1999;	99US-0152363P.
PR	28-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99US-0153070P.
PR	29-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	30-JUN-1999;	99US-0141287P.	PR	15-SEP-1999;	99US-0154018P.
PR	01-JUL-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.
PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144085P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144332P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160814P.
PR	23-JUL-1999;	99US-0145145P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160981P.
PR	26-JUL-1999;	99US-0145276P.	PR	22-OCT-1999;	99US-0160989P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145919P.	PR	25-OCT-1999;	99US-0161406P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161359P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146388P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161920P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161992P.
PR	04-AUG-1999;	99US-0147204P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147302P.	PR	29-OCT-1999;	99US-0162142P.
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			

Query Match

Best Local Similarity

Mismatches

Conservative

Mismatches

Indels

Gaps

7;

Qy	1	MALAFSTTSISAPTSTPPSS----	BKLPQIGSIRISERPIGCAVNFNLS-QRRSLVKPV	54
Db	23	MAWVNVSSSSSGIINSRFGVSLKPKVQISRLRLDRVHVAPVSLNLSCKSSSVKPL	82	
Qy	55	NAEPPRKDSIVPLAATTIVASATKE---	EDPEQIASLDNAPSLIEMDRALDKFGNDI	111
Db	83	NAEPTKDSMIPLAATWVAIEAEVEVEIEDFEELAKKLENASPLEIMDKALEKYGNDI	142	
Qy	112	ATAFSGAEDVALIEYAKLTGRPFVPSLDTGRINPETYQLFDAVEKHGIRIYMPPDV	171	
Db	143	ATAFSGAEDVALIEYAKLTGRPFVPSLDTGRINPETYQLFDAVEKHGIRIYMPPDV	202	
Qy	172	EQQALVRSKGLFSFYEDGHQCCRVKVRPLRRALKGLRAWITGQRKDSQPGTRSEIPVV	231	
Db	203	EQQGLVRSKGLFSFYEDGHQCCRVKVRPLRRALKGLRAWITGQRKDSQPGTRSEIPVV	262	
Qy	232	QVDPAPFEGMDGGIGSLVKWNPVANVKGHDINWPLRTMNVPNVNSLHAKGYVSIQCEPCTRP	291	

Db 263 QVDVPVREGLDGGVSLVKMNPVANVEGNDVMNFLTMDVPNTLHAAAGYISIGCEPCTKA 322  
Qy 292 VLFQOHEREGRWWEDAKAECGLHKGNVKQOKEEDVNGLSQSHANGDATTTVPDIFNS 351  
Db 323 VLFQOHEREGRWWEDAKAECGLHKGNVK-ENSDDAKVNGESKS-----AVADIFKS 374  
Qy 352 PNVVNLSTGTIENLAKLEDKRPWLVLVYAPWCPYCOAMBEESYVDLADKLAGSTGMKVKG 411  
Db 375 ENLVTLRSQIENLMLKENRKEPIWVLVYAPWCPFCQAMASVDELADKLAGS-GIKVAK 433  
Qy 412 FRADGEQKEFAKSELQSGFPTTLFFPKHSSRPTIKYPSEKRDVDSLMAFVNALR 466  
Db 434 FRADGDQKEFAKSELQSGFPTTLVFPKNGSRP-IKYPSEKRDVESLTSFLNLVR 487  
  
RESULT 8  
ID AAG42685  
AC AAG42685 standard; protein; 463 AA.  
XX  
XX AAG42685;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 53261.  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
XX hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
XX  
XX Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-00301439.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
XX 08-MAR-1999; 99US-0123180P.  
XX 09-MAR-1999; 99US-0123548P.  
XX 23-MAR-1999; 99US-0125788P.  
XX 25-MAR-1999; 99US-0126264P.  
XX 29-MAR-1999; 99US-0126785P.  
XX 01-APR-1999; 99US-0127462P.  
XX 06-APR-1999; 99US-0128234P.  
XX 08-APR-1999; 99US-0128714P.  
XX 16-APR-1999; 99US-0129845P.  
XX 19-APR-1999; 99US-0130077P.  
XX 21-APR-1999; 99US-0130449P.  
XX 23-APR-1999; 99US-0130510P.  
XX 23-APR-1999; 99US-0130891P.  
XX 28-APR-1999; 99US-0131449P.  
XX 30-APR-1999; 99US-0132048P.  
XX 30-APR-1999; 99US-0132407P.  
XX 04-MAY-1999; 99US-0132484P.  
XX 05-MAY-1999; 99US-0132485P.  
XX 06-MAY-1999; 99US-0132486P.  
XX 06-MAY-1999; 99US-0132487P.  
XX 07-MAY-1999; 99US-0132863P.  
XX 11-MAY-1999; 99US-0134256P.  
XX 14-MAY-1999; 99US-0134218P.  
XX 14-MAY-1999; 99US-0134219P.  
XX 14-MAY-1999; 99US-0134221P.  
XX 14-MAY-1999; 99US-0134370P.  
XX 18-MAY-1999; 99US-0134768P.  
XX 19-MAY-1999; 99US-0134941P.  
XX 20-MAY-1999; 99US-0135124P.  
XX 21-MAY-1999; 99US-0135353P.  
XX 24-MAY-1999; 99US-0135629P.  
XX 25-MAY-1999; 99US-0136021P.  
XX 27-MAY-1999; 99US-0136392P.  
XX 28-MAY-1999; 99US-0136782P.  
XX 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 23-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.



PR 30-APR-1999; 99US-01332048P.  
PR 30-APR-1999; 99US-01332407P.  
PR 04-MAY-1999; 99US-01322484P.  
PR 05-MAY-1999; 99US-01332485P.  
PR 06-MAY-1999; 99US-01332486P.  
PR 06-MAY-1999; 99US-01332487P.  
PR 07-MAY-1999; 99US-01332863P.  
PR 11-MAY-1999; 99US-01334256P.  
PR 14-MAY-1999; 99US-01334218P.  
PR 14-MAY-1999; 99US-01334219P.  
PR 14-MAY-1999; 99US-01334221P.  
PR 14-MAY-1999; 99US-01334370P.  
PR 18-MAY-1999; 99US-01334768P.  
PR 18-MAY-1999; 99US-01334941P.  
PR 20-MAY-1999; 99US-01335124P.  
PR 20-MAY-1999; 99US-01335353P.  
PR 21-MAY-1999; 99US-01335629P.  
PR 24-MAY-1999; 99US-01336021P.  
PR 25-MAY-1999; 99US-01336392P.  
PR 27-MAY-1999; 99US-01336782P.  
PR 28-MAY-1999; 99US-01337222P.  
PR 01-JUN-1999; 99US-01337528P.  
PR 03-JUN-1999; 99US-01337502P.  
PR 04-JUN-1999; 99US-01337724P.  
PR 07-JUN-1999; 99US-01338034P.  
PR 08-JUN-1999; 99US-01338540P.  
PR 10-JUN-1999; 99US-01338847P.  
PR 10-JUN-1999; 99US-01339119P.  
PR 14-JUN-1999; 99US-01339452P.  
PR 16-JUN-1999; 99US-01339453P.  
PR 16-JUN-1999; 99US-01339482P.  
PR 17-JUN-1999; 99US-01339454P.  
PR 18-JUN-1999; 99US-01339455P.  
PR 18-JUN-1999; 99US-01339456P.  
PR 18-JUN-1999; 99US-01339457P.  
PR 18-JUN-1999; 99US-01339458P.  
PR 18-JUN-1999; 99US-01339461P.  
PR 18-JUN-1999; 99US-01339462P.  
PR 18-JUN-1999; 99US-01339463P.  
PR 18-JUN-1999; 99US-01339750P.  
PR 18-JUN-1999; 99US-01339763P.  
PR 21-JUN-1999; 99US-01339817P.  
PR 22-JUN-1999; 99US-01339899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 24-JUN-1999; 99US-0140823P.  
PR 28-JUN-1999; 99US-0140931P.  
PR 29-JUN-1999; 99US-0141287P.  
PR 30-JUN-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 01-JUL-1999; 99US-0142055P.  
PR 02-JUL-1999; 99US-0142330P.  
PR 06-JUL-1999; 99US-0142803P.  
PR 08-JUL-1999; 99US-0142920P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 16-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.

PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 26-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.





```
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144844P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145152P.
PR 22-JUL-1999; 99US-0145159P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.

PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 73.9%; Score 1806.5; DB 3; Length 487;
Best Local Similarity 73.7%; Pred. No. 4.2e-177;
Matches 350; Conservative 48; Mismatches 58; Indels 19; Gaps 7;

Qy 1 MALAFTSSISAPTSTFPSS-----EPKLPQIGSIRISERPITGGAVNFNLS-ORSLVKPV 54
Db 23 MAMSVNVSSSSSGIINSRFGVSLPEKVSQIGSLRLDLRVHVPVSLNLGKRSSSVKPL 82
Qy 55 NAEPPRKDSIVPLAATTIVASASETKE---EDPEQIASDLDNASPLEIMDRALDKFGNDI 111
Db 83 NAEPKTNDSMIPLAANTWVAIEAEVVEVEEDFELAKLENASPLEIMDKALEKYGNDI 142
Qy 112 AIAFSGAEDVALIEYAKLTGRPPRVFSLDTGRNLNPETYQLFDAVEKHYGIRIETMFPDVA 171
Db 143 AIAFSGAEDVALIEYAKLTGRPPRVFSLDTGRNLNPETYQLFDAVEKHYGIRIETMFPDVS 202
Qy 172 EVQALVRSGKLFSEYEDGHOECRCRVRKVRPLRRALKGLRAWITGQRKDOOSPGRSEIPVV 231
Db 203 EVQGLVRSGKLFSEYEDGHOECRCRVRKVRPLRRALKGLKAWITGQRKDOOSPGRSEIPVV 262
Qy 232 QVDPAFEGMDGGIGSLVKWNPVANVKGHDITWNPRTWNPVNSLHAKGYVSGICEPCTR 291
Db 263 QVDPVFGLDGGVSLVKWNPVANVEGNDVWNPRTWNPVNTLHAAGYISIGCEPCTKA 322
Qy 292 VLPQHEREGRWWEWEDAKAKECGLHKGNVKQKKEEDVNGNGLSQSHANGDATTVPDIFNS 351
Db 323 VLPQHEREGRWWEWEDAKAKECGLHKGNVK-ENSDDAKVNGESKS-----AVADIFKS 374
Qy 352 PNVNLSRTGIENLAKLEDKPEWLVLYAPWCYQCAWEEYVDLADKLAGSTGMKVKG 411
Db 375 ENLVTLRQGIENLMLKLENKPEWLVLYAPWCYQCAWEEYVDLADKLAGS-GIKVAK 433
```





PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 21-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 23-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142820P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143342P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 23-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 26-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147915P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.

PR	25-OCT-1999;	99US-0161405P.	PR	25-MAR-1999;	99US-0126264P.
PR	25-OCT-1999;	99US-0161406P.	PR	25-MAR-1999;	99US-0126785P.
PR	26-OCT-1999;	99US-0161359P.	PR	01-APR-1999;	99US-0127462P.
PR	26-OCT-1999;	99US-0161360P.	PR	06-APR-1999;	99US-0128234P.
PR	26-OCT-1999;	99US-0161361P.	PR	08-APR-1999;	99US-0128714P.
PR	28-OCT-1999;	99US-0161920P.	PR	16-APR-1999;	99US-0129845P.
PR	28-OCT-1999;	99US-0161922P.	PR	19-APR-1999;	99US-0130077P.
PR	28-OCT-1999;	99US-0161933P.	PR	21-APR-1999;	99US-0130449P.
PR	28-OCT-1999;	99US-0162142P.	PR	23-APR-1999;	99US-0130510P.
PR	28-OCT-1999;	99US-0162142P.	PR	23-APR-1999;	99US-0130891P.
Query Match 73.7%; Score 1801; DB 3; Length 458;					
Best Local Similarity 74.0%; Pred. No. 1.4e-176;					
Matches 348; Conservative 51; Mismatches 55; Indels 16; Gaps 6;					
Qy	1	MALAF-----TSSISAPTSPSSPEKLPQIGSRISRISRPIGGAVFNLSQRRSLVKPVNA	56		
Db	1	MALAINVSSSSSAISSPSSDLKVTKIGSLRLNRTNVSASLSLGRSSVKALNV	60		
Qy	57	EPPRKDSIVPLAATTIVASASETKEEDPEOIASOLDNASPLEIMDRALDKFGNDIAIAPS	116		
Db	61	QSITKESIV--ASEVTEKLDVVEDEFEELAKLENASPLEIMDKALEFGNDIAIAPS	117		
Qy	117	GAEDVALIEYAKLTGRPRVFSLDTGRINPETYQLFDAVEKHGIRIEYMPDPAVEYOAL	176		
Db	118	GAEDVALIEYAKLTGRPRVFSLDTGRINPETYQLFDAVEKHGIRIEYMPDPAVEYOAL	177		
Qy	177	VRSGKLSFVEDGHOECRCVRKVRPIRRALKGLRAWITGORKDQSPGTRSEIPVQVDP	236		
Db	178	VRNKGKLSFVEDGHOECRCIRKVRPLRRALKGLRAWITGORKDQSPGTRSEIPVQVDP	237		
Qy	237	PEGMDGGISIVKKNPVANVKKGDIWNFLRTMNPVNSLHAKGVYSIGCEPCTRPVLPQ	296		
Db	238	PEGLDGGVSLVKNPVANVEGDNVNFRTMDVPVNLHAGVYSIGCEPCTRAVLPGQ	297		
Qy	297	HEREGRWWDADAKACEGLHKNVKKQKEEDVNGSLSQSHANGDATTVDPIDFNSPVNV	356		
Db	298	HEREGRWWDADAKACEGLHKNVKKQKEEDVNGSLSQSHANGDATTVDPIDFNSPVNV	350		
Qy	357	LSRTGIENLAKLEDKRPWLVLVAPWCPCQAMERSYVDLADKLAGSTGMKVKCFRADG	416		
Db	351	LSRQGIENLAKLENKREAWIVLVAPWCPCQAMEASFDELADKLGS-GYKVAKFRADG	409		
Qy	417	EQKEFAKSELQGSFPITLPPKHSSRPTIKYPSKEDVDLSMAFVNALR	466		
Db	410	DQKDFAKKELQGSFPITLVFPKNSRP-IKYPSEKDDVSLTSFLNLVR	458		
RESULT 13					
ID	AAG20362	standard; protein; 483 AA.			
XX	AC	AAG20362;			
XX	DT	17-OCT-2000 (first entry)			
XX	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 22519.			
XX	KW	Protein identification; signal transduction pathway; metabolic pathway;			
XX	KW	hybridisation assay; genetic mapping; gene expression control; promoter;			
XX	OS	termination sequence.			
XX	FN	Arabidopsis thaliana.			
XX	PD	EPI033405-A2.			
XX	FF	06-SEP-2000.			
XX	FF	25-FEB-2000; 2000EP-00301439.			
PR	25-FEB-1999;	99US-0121825P.			
PR	05-MAR-1999;	99US-0123180P.			
PR	09-MAR-1999;	99US-0123548P.			
PR	23-MAR-1999;	99US-0125788P.			

PR	25-MAR-1999;	99US-0126264P.
PR	25-MAR-1999;	99US-0126785P.
PR	01-APR-1999;	99US-0127462P.
PR	06-APR-1999;	99US-0128234P.
PR	08-APR-1999;	99US-0128714P.
PR	16-APR-1999;	99US-0129845P.
PR	19-APR-1999;	99US-0130077P.
PR	21-APR-1999;	99US-0130449P.
PR	23-APR-1999;	99US-0130510P.
PR	23-APR-1999;	99US-0130891P.
PR	28-APR-1999;	99US-0131449P.
PR	30-APR-1999;	99US-0132048P.
PR	30-APR-1999;	99US-0132407P.
PR	04-MAY-1999;	99US-0132484P.
PR	05-MAY-1999;	99US-0132485P.
PR	06-MAY-1999;	99US-0132486P.
PR	06-MAY-1999;	99US-0132487P.
PR	11-MAY-1999;	99US-0132863P.
PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	14-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.









PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142309P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144008P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145227P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151080P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Search completed: February 18, 2006, 14:33:18  
Job time : 190 secs

**THIS PAGE LEFT BLANK**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2006, 14:33:35 ; Search time 40 Seconds  
(without alignments)  
1120.925 Million cell updates/sec

Title: US-10-731-525-8

Perfect score: 2445

Sequence: 1 MALAFTSSISAPSTPPSS.....KYPSEKRDVDSLMAFVNALR 466

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_80:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1811.5	74.1	465	2 B85058	5'-adenyllylsulfate
2	1797	73.5	458	2 T49106	PRH26 protein - Ar
3	1788.5	73.1	463	2 T10065	phosphoadenyllyl-su
4	1738	71.1	454	2 C96648	hypothetical prote
5	1648	67.4	406	2 S71242	phosphoadenyllyl-su
6	1555.5	63.6	422	2 S71243	phosphoadenyllyl-su
7	1216	49.7	423	2 T52251	5'-adenyllylsulfate
8	745.5	30.5	267	2 H83426	3'-phosphadenosin
9	519.5	21.2	239	2 D90470	hypothetical prote
10	470.5	19.2	244	2 H81905	probable phosphoad
11	463	18.9	246	2 C81111	phosphadenosine p
12	406	16.6	260	2 H97458	hypothetical prote
13	406	16.6	260	2 AB2677	phosphadenosine p
14	325	13.3	196	2 C69839	3'-phosphadenosin
15	305.5	12.5	254	2 C70682	probable cybH - My
16	302.5	12.4	231	2 F83835	3'-phosphadenosin
17	296.5	12.1	233	2 H69611	phosphadenyllyl-su
18	290.5	11.9	244	2 H97388	phospho-adenyllysu
19	268.5	11.0	252	2 AG3472	phosphadenyllyl-su
20	259	10.6	241	2 AH2363	phosphadenosine p
21	244.5	10.0	255	2 C75594	phosphadenosine p
22	235.5	9.6	237	2 E82674	3'-phosphadenosin
23	228	9.3	239	1 S34193	phosphadenyllyl-su
24	214	8.8	266	2 S67435	phosphadenyllyl-su
25	213	8.7	244	2 AE0858	3'-phosphadenosin
26	211	8.6	244	1 C34354	phosphadenyllyl-su
27	210	8.6	244	1 RDECPA	phosphadenyllyl-su
28	210	8.6	261	2 S59826	phosphadenyllyl-su
29	207	8.5	244	2 B85926	3'-phosphadenosin

30	207	8.5	244	2 A91081	3'-phosphadenosin
31	197	8.1	253	2 D82329	phosphadenosine p
32	193.5	7.9	249	2 S74657	phosphadenyllyl-su
33	168.5	6.9	244	2 D84979	phosphadenosine p
34	167.5	6.9	632	2 B75215	5'-adenyllylphospho
35	162.5	6.6	232	1 S28609	phosphoadenyllyl-su
36	157.5	6.4	480	2 B64308	hypothetical prote
37	154.5	6.3	266	2 C96913	sulfate adenylate
38	154	6.3	532	2 T07927	protein disulfide-
39	154	6.3	638	1 ISMSER	protein disulfide-
40	154	6.3	645	1 A23723	protein disulfide-
41	151	6.2	630	2 E71451	hypothetical prote
42	146.5	6.0	443	2 D86183	hypothetical prote
43	145	5.9	643	1 S32476	protein disulfide-
44	143	5.8	504	2 S41661	protein disulfide-
45	141	5.8	411	2 E64421	hypothetical prote

ALIGNMENTS

RESULT 1

B85058

5'-adenyllylsulfate reductase [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004

C;Accession: B85058

R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sprin

Nature 402, 769-777, 1999

A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A;Reference number: A85001; MUID:20083488; PMID:10617198

A;Accession: B85058

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-465 <STO>

A;Cross-references: UNIPROT:P92979; UNIPARC:UPI000000BD8A; GB:NC\_001268; NID:g7267219; P

C;Genetics:

A;Gene: AT4g04610

A;Map position: 4

Query Match 74.1%; Score 1811.5; DB 2; Length 465;  
Best Local Similarity 73.9%; Pred. No. 1.7e-132;  
Matches 351; Conservative 48; Mismatches 57; Indels 19; Gaps 7;

QY	1	MALAFTSSISAPSTPPSS-----EPKLPQIGSIRISERPIGAVNFNLS-QRRSLVKPV	54
DB	1	MMSVNVSSSSSGIINSRFGVLEPKVQIGSLRLDRVHVAVPVSLSNLSKSSSSVKPL	60
QY	55	NAEPPRKDSIVPLAATTIVASASETKE---EDFEQIASLDLNASPLEIMDRALDKFGNDI	111
DB	61	NAEPTKDSMIPLAATMWAEIAEEVEVEIEDEFEELAKKLENASPLEIMDKALEKYNDI	120
QY	112	ATAFSGAEDVALIEYAKLTGRPPRVPSLDTGRINLPETYQLFDAVEKHGIRIIRIYMPDPAV	171
DB	121	ATAFSGAEDVALIEYAKLTGRPPRVPSLDTGRINLPETYRFFDAVEKHGIRIIRIYMPDVS	180
QY	172	EQVALVRSKGLSFYEDGHQECRCVRKVRPLRALGLRAWITGQRKQSPGTRSRIPVV	231
DB	181	EQVGLVRSKGLSFYEDGHQECRCVRKVRPLRRALKGLKAWITGQRKQSPGTRSRIPVV	240
QY	232	QVDPAPEGMGGTSGSLVKKNPVANVKGHDIMFLRTMNPVNSLHAKGVYVSGICEPCTRP	291
DB	241	QVDVPFEGLDGGVSLVKKNPVANVKGHDIMFLRTMNPVNSLHAKGVYVSGICEPCTKA	300
QY	292	VLPQGHERRGWWEDEAKKCGCLHKGNYKQKKEEDVNGSLSQSHANGDATTVPDIFNS	351
DB	301	VLPQGHERRGWWEDEAKKCGCLHKGNYK-ENSDDAKVNAGESK-----AVADIFKS	352
QY	352	PNVNLSPRTGNIENLAKLEDRKEPWLVLVYAPCPYCOAMEESVVDLADKLAGSTGMKVGK	411
DB	353	ENLVTLTSGQGIENMLKLENKBPWIVLVYAPWCPFCOAMEASYDELDADKLAGS-GIKVAK	411
QY	412	FRADGQKEFAKSELQSGSFPTILFPFKGHSRRPTIKYPSEKRDVDSLMAFVNALR	466







RESULT 7  
T52251  
5'-adenyllysulfate reductase [EC 1.8.99.-] [validated] - green alga (Enteromorpha intest  
C;Species: Enteromorpha intestinalis (hollow green seaweed)  
C;Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 09-Jul-2004  
C;Accession: T52251  
R;Gao, Y.; Schofield, O.M.; Leustek, T.  
Plant Physiol. 123, 1087-1096, 2000  
A;Title: Characterization of sulfate assimilation in marine algae focusing on the enzyme  
A;Reference number: Z26003; MUID:20349720; PMID:10889258  
A;Accession: T52251  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-423 <GAO>  
A;Cross-references: UNIPROT:O81350; UNIPARC:UPI000000A273A; EMBL:AF069951; PIDN:AAAC26855.  
C;Function:  
A;Description: catalyzes reduced glutathione-dependent reduction of adenylylsulfate to a  
C;Keywords: oxidoreductase

Query Match 49.7%; Score 1216; DB 2; Length 423;  
Best Local Similarity 60.1%; Pred. No. 2.1e-86;  
Matches 239; Conservative 42; Mismatches 99; Indels 18; Gaps 6;  
C;Superfamily: 3'-phosphoadenosine 5'-phosphosulfate reductase

Qy 68 AATTIVASSETKEEDFEQIASDLNDSAPLEIMDRALDKGNDIAIAPSGAEDVALIEYA 127  
Db 40 AAAAATAAPEGAKATDAAETKELNNKSPLEIMDHATFGDEVAIAIAPSGAEDVALIEYA 99  
Qy 128 KLGRPRFVSLDTGRNLNPETYQLFDAVEKHGIRIEMFPDPAVEQALVRSKGLFSFYE 187  
Db 100 HLGRKYRVSLDTGRNLNPETYQLFDAVEKHGIRIETFPDAQETDVLREKGLFSFYE 159  
Qy 188 DGHQECRRVKRPLRALKGLRAWITGQRKDQSPGTRSEIPVQVDPDAFEGMDGGIGSL 247  
Db 160 DGHQECRRVKRPLRLQLKLTQAWITGQRKDQSPGTRMEVPAVQVDPFEGQSGDGL 219  
Qy 248 VKNPNVANVKGHDITNFLTMMNPVNSLHAKGYVSGCEPCTRPVLPQHEREGRWMMEDAKA 307  
Db 220 VKYNPLNTMTSAEYNNFLRMNVPNSKLHCEGYVSGCEPCTRPVLPNQEREGRWMMED 279  
Qy 308 AKAKECGLHKNVKKQKEEDVNGLSQSHANGDATTVPDIFNSPNVNLSTRTGIENLAK 367  
Db 280 SAAKECGLHSGNVDSAEAKAE-----APDLWTGCGAVALDKATLEKLAG 326  
Qy 368 LEDRKEPWLVLVYAPWCPYCOAMBEESYVDLADKLAGSTGMKVGFRADGEOKEPAKSELQ 427  
Db 327 -GDRDKDTMVLVYAPWCPFCQAFEPDYEFPVAQE-PGGAGFGVAKYNADAD-REYSES-LG 382  
Qy 428 LGSPTTILFFPKHSSRTIKYPSEKRDVDSLMAFVNAL 465  
Db 383 LKTPTTIFLPKGS DK-VVKPESDRRTVESMTWMAAAL 419

RESULT 8  
H83426  
3'-phosphoadenosine-5'-phosphosulfate reductase PA1756 [imported] - Pseudomonas aerugin  
C;Species: Pseudomonas aeruginosa  
C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C;Accession: H83426  
R;Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A;Reference number: A82950; MUID:20437337; PMID:10984043  
A;Accession: H83426  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-267 <STO>  
A;Cross-references: UNIPROT:O05927; UNIPARC:UPI0000128CED; GB:AE004601; GB:AE004091; NID  
A;Experimental source: strain PA01  
C;Genetics:  
A;Gene: cyhH; PA1756

Query Match 30.5%; Score 745.5; DB 2; Length 267;  
Best Local Similarity 54.8%; Pred. No. 3.2e-50;  
Matches 130; Conservative 46; Mismatches 59; Indels 3; Gaps 1;  
C;Superfamily: 3'-phosphoadenosine 5'-phosphosulfate reductase

Qy 83 DFEQIASDLNDSAPLEIMDRALDKGNDIAIAPSGAEDVALIEYAKLTGRPFRVFSLDTG 142  
Db 29 DLPALASSLADKSPQDILKAAPFHFGEDELWISFGAEDVVLVDMAWLNKRVKFSLDTG 88  
Qy 143 RLNPETYQLFDAVEKHGIRIEMFPDPAVEQALVRSKGLFSFVEDGHQECRRVKRPL 202  
Db 89 RLHPETVRFIDQVREHYGIAIDVLSPPRLLEPLVKEKGLFSFYRDGHGECGGRKIEPL 148  
Qy 203 RRALKGLRAWITGQRKDQSPGTRSEIPVQVDPDAFEGMDGGIGSLVKNPNVANVKGHDII 262  
Db 149 KRLAGVRAWATGQRDQSGTSTQSAVLBIIDGAFSTPE---KPLYKFNPLSSMTSBEVM 205  
Qy 263 NFLRTMMNPVNSLHAKGYVSGCEPCTRPVLPQHEREGRWMMEDAKA 320  
Db 206 GYIRMLELPYNSLHERGYISIGCEPCTRPVLPQHEREGRWMMEDAKA 263

RESULT 9  
D90470  
hypothetical protein cyhH [imported] - Sulfolobus solfataricus  
C;Species: Sulfolobus solfataricus  
C;Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004  
C;Accession: D90470  
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-  
Jong, I.; Jeffries, A.C.; Kozera, C.W.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.  
arrett, R.A.; Ragan, M.A.; Sensesen, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A;Description: Sulfolobus solfataricus complete genome.  
A;Reference number: A99139  
A;Accession: D90470  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-239 <KUR>  
A;Cross-references: UNIPROT:Q97UT3; UNIPARC:UPI000000648C0; GB:AE006641; NID:gl3816282; P  
C;Genetics:  
A;Gene: cyhH  
C;Superfamily: 3'-phosphoadenosine 5'-phosphosulfate reductase

Query Match 21.2%; Score 519.5; DB 2; Length 239;  
Best Local Similarity 43.3%; Pred. No. 8.8e-33;  
Matches 106; Conservative 48; Mismatches 78; Indels 13; Gaps 7;  
C;Superfamily: 3'-phosphoadenosine 5'-phosphosulfate reductase

Qy 81 BEDFEQIASDLNDSAPLEIMDRALDKGNDIAIAPSGAEDVALIE-YAKLTGRPFRVFS 138  
Db 4 KDDELSTLKLFLFEDKEPLEVLKWKIEKYPKIALACSLQAEEDLVILDMLSKYTEKP-RVFI 62  
Qy 139 LDTGRNLNPETYQLFDAVEKHGIRIEMFPDPAVEQALVRSKGLFSFVE--DGHQECRRV 196  
Db 63 IDTGRHQESYDLTEEVKYNKYNIDLRIVFPDYKEVEDLVNKNYGINFLFKSVELRKACCEV 122  
Qy 197 RKVPLRLALKGLRAWITGQRKDQSPGTRSEIPVQVDPDAFEGMDGGIGSLVKNPNVANV 256  
Db 123 RKVPLRLALDKGMAWITGURRQEQ-FTGRGRKRIEIDE----VNGGI---IKUNPLADW 174  
Qy 257 KGHDIWNLFRMTNVPVNSLHAKGYVSGCEPCTRPVLPQHEREGRWMMEDAKA 316  
Db 175 TWEQWVEYKNNVYNNKLYDKGYSIGCVPTCPVKEPWEHPRAGRMMWQNSDKECGLH 234  
Qy 317 KGNVYK 321  
Db 235 YREVK 239

RESULT 10  
H81905  
probable phosphoadenosine phosphosulfate reductase NMA1366 [imported] - Neisseria mening  
C;Species: Neisseria meningitidis

```

Qy 49 SLVKPVNAEPPRKDSIVPLAATTIVASATKEEDFEQIASDLNDNASPLEIMDRALDKFG 108
Db 4 TLFKPALWQIPIHIGS-----GGETAIAKTEKTTKQRLHRIVGSRDA-----RFA 48
Qy 109 NDIAIAFSGAEDVALLEYAKLTGRPPRVFSLDTGRNLNPEYQLFDVAEKHY-GIRIEYMF 167
Db 49 SSIA-----AEDWITDLIAGENLNTGIFLDTGLLHTETLNLDDRGRAYPHLRIKFR 103
Qy 168 PDAVEQALVRSKGLFSFYB--DGHQCCRRVKRVLRLRAKGLRAWITQQRKDQSGTR 225
Db 104 PVREDADRYVESKGRFAFYDSVEARRECCRIRKTEPLNRAIAGADAWLTGQREQS-ATR 162
Qy 226 SEIPVQVQDPAPFGMDGGIGSLVKWNPVANVKGHDWFLRTWNPVNSLHAKYVSGC 285
Db 163 TELPFAEYDAG-----RGIG--KYNPIDFOWSHDVWAYILANNVPYNDLYRQGFPSIGC 214
Qy 286 EPTCRPVLPOCHEREGRWMDAKAKCEGLHK 317
Db 215 DPCTRPVKAGEDIRAGRWWEGRNSKEGLHK 246
RESULT 12
H97458
hypothetical protein AGR_C_1497 [imported] - Agrobacterium tumefaciens (strain
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: H97458
A.; Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.;
R.; Liu, P.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Mar
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobact
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: H97458
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-260 <KUR>
A;Cross-references: UNIPROT:Q8UH67; UNIPARC:UPI0000128CE6; GB:AE007869; PIDN:AA
C;Genetics: C
A;Gene: AGR_C_1497
A;Map position: circular chromosome
C;Superfamily: 3'-phosphoadenosine 5'-phosphosulfate reductase
Query Match 16.6%; Score 406; DB 2; Length 260;
Best Local Similarity 35.4%; Pred.No.6.2e-24;
Matches 92; Conservative 37; Mismatches 117; Indels 14; Gaps 4;
Qy 65 VPLAATTIVASATKEEDFEQIASDLNDNASPLEIMDRALDKFGNDIAIAFSGAEDVALI 124
Db 3 IPDVTMTINSTNASADTASLDATLAGLDLAGLSFV---AGLGRAVFTTSLGIEDQVIT 59
Qy 125 EYAKLTGRPPRVFSLDTGRNLNPEYQLFDVAEKHYGIRIEYMPDPDAVEQALVRSKGLFS 184
Db 60 AALGTHRLPTDVTLETGRLFKEVTDLIDETERFGEIRRFRRPEQDDIDAYAAYKGLNG 119
Qy 185 FYB--DGHQCCRRVKRVLRLRAKGLRAWITQQRKDQSGTRSEIPVQVDPAPFGMDG 242
Db 120 FYGSEARHACCHVRKLIPLGKALEGAATWITGLRQGS--GNRAATPFPAEFDAE----- 172
Qy 243 GIGSLVKWNPVANVKGHDWFLRTWNPVNSLHAKYVSGICEPCTRPVLPQCHEREGR 302
Db 173 --RNLIKINALADWDIEQIRAYVAENIPWNPVPLHQRGYPSIGCEPCTRAIKPCEPERAGR 230
Qy 303 WWDADAKACEGLHKGNVKQ 322
Db 231 WWDENDEKCEGLHVAGAEQ 250
RESULT 13
AB2677
phosphoadenosine phosphosulfate reductase [imported] - Agrobacterium tumefaciens
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004

```



Job time : 41 secs

**THIS PAGE LEFT BLANK**





```

QY 171 VEVOALVRSKGLRFSYEDGHOECRRVVRPLRRALKGLRAWITGQRKDSQPGTRSEIPV 230
DB 180 VEVOALVRTKGLRFSYEDGHOECRRVVRPLRRALKGLRAWITGQRKDSQPGTRSEIPV 239
QY 231 VQVDPAPFEGDGGIGSLVKNPVPANVKGHDIMNPLRTMNVVNSLHAKGVYSIGCEPCTR 290
DB 240 VQVDPVFEGLDGGIGSLVKNPVPANVGLDWSFLRTMDVPVNSLHSGQVYSIGCEPCTR 299
QY 291 PVLPGQHERGRWWEDAKAKCEGLHKNVQKQKEEDVNGLSQSHANGDATTPVDIFN 350
DB 300 PVLPGQHERGRWWEDAKAKCEGLHKNVQKQKEEDVNGLSQSHANGDATTPVDIFN 356
QY 351 SPNVNLSRTGIENLAKLEDRKEPWLVLVYAPWCPYQOAMEESYVDLADKLKAGSTGMKV 410
DB 357 SQDVVLSRSRTGIENLAKLEDRKEPWLVLVYAPWCPYQOAMEESYVDLADKLKAGS-GYKVA 415
QY 411 KFRADGQKEFAKSELQGSFPTILFPFKHSSRPTIKYPSEKRDVDSLMAFVNALR 466
DB 416 KFRADGQKEFAKSELQGSFPTILFPFKHSSQP-IKYPSEKRDVDSLMAFVNALR 470

RESULT 2
Q6V3B0_9ROSI PRELIMINARY; PRT; 465 AA.
AC Q6V3B0;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DE Adenosine 5' phosphosulfate reductase.
OS Populus alba x Populus tremula.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Malpighiales; Salicaceae; Salicaceae; Populus.
OX NCBI_TaxID=80863;
RN NUCLEOTIDE SEQUENCE.
RP Kopriva S., Hartmann T., Masearo G., Hoenicke P., Rennenberg H.;
RA "Regulation of sulfate assimilation by nitrogen and sulfur nutrition
RT in poplar trees.";
RL Trees 18:320-326 (2004) .
RN NUCLEOTIDE SEQUENCE.
RX PubMed=15020643; DOI=10.1093/ixb/erh094;
RA Hartmann T., Honicke P., Wirtz M., Hell R., Rennenberg H., Kopriva S.;
RT "Regulation of sulphate assimilation by glutathione in poplars
RT (Populus tremula x P. alba) of wild type and overexpressing gamma-
RT glutamylcysteine synthetase in the cytosol.";
RL J. Exp. Bot. 55:837-845 (2004) .
DR EMBL; AY353089; AAQ57202.1; -; mRNA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0016671; F:oxidoreductase activity, acting on sulfur g. . .; IEA.
DR GO; GO:0016740; P:transferrase activity; IEA.
DR GO; GO:0008652; P:amino acid biosynthesis; IEA.
DR GO; GO:0019344; P:cysteine biosynthesis; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0019421; P:sulfate reduction, APS pathway; IEA.
DR InterPro; IPR004508; APS_reduc.
DR InterPro; IPR011798; APS_reductase.
DR InterPro; IPR002500; PAPS_reduct.
DR InterPro; IPR006662; ThioRedox.
DR InterPro; IPR006663; ThioRedox_dom2.
DR Pfam; PF01507; PAPS_reduct; 1.
DR PRINTS; PR00421; THIOREDOXIN.
DR TIGRFAMs; TIGR00424; APS_reduc; 1.
DR TIGRFAMs; TIGR02055; APS_reductase; 1.
SQ SEQUENCE 465 AA; 51540 MW; 19A9B6A4C934AC21 CRC64;

```

Query Match 77.0%; Score 1883.5; DB 2; Length 465;  
 Best Local Similarity 77.9%; Pred. NO. 1.7e-141;  
 Matches 366; Conservative 35; Mismatches 60; Indels 9; Gaps 7;

```

QY 1 MALAFTSSISAPTSTF--PSSEPKLPQIGSIRISERPIGGAIVNPLSQRRSLVKVPVNAEP 58
DB 1 MALSSSSSISAGSFRSSQELKAPQFGSKLLDRT-TFVNVS--SQRCVAKPLNAEP 58
QY 59 PRKDSIVPLAATTIVASASE--TKBEDFQIASDLDNASPLEIMDRALDFKGDIAIAFS 116
DB 59 KRNGSVVPLAATTAAPEIAEKVEVEDEYELAKELVNGSPLEIMDKALEKFGDDIAIAFS 118
QY 117 GAEDVALIEYAKLTGPRFRVPSLDTGRLNPETTPQLPDAVEKHGIRIERYMFPDVAEQAL 176
DB 119 GAEDVALIEYAKLTGPRFRVPSLDTGRLNPETTHFFDQVEKHGIRIERYMFPDVAEQAL 178
QY 177 VRSKGLFSFYEDGHOECRRVVRPLRRALKGLRAWITGQRKDSQPGTRSEIPVQVDP 236
DB 179 VRNKGLSFYEDGHOECRRVVRPLRRALKGLRAWITGQRKDSQPGTRSEIPVQVDPV 238
QY 237 FEGDGGIGSLVKNPVPANVKGHDIMNPLRTMNVVNSLHAKGVYSIGCEPCTRPLVPGQ 296
DB 239 FEGLDGCGAGSLIKWNPMPANVEGDQVWKFLRTMDVPVNSLHSGYISIGCEPCTRPLVPGQ 298
QY 297 HEREGRWWEWEDAKAKCEGLHKNVQKQKEEDVNGLSQSHANGDATTPVDIFNSPNVN 356
DB 299 HEREGRWWEWEDAKAKCEGLHKNLQKQDAQLNGNGAAHANG-AATVADIFENSENLVN 357
QY 357 LSRTGIENLAKLEDRKEPWLVLVYAPWCPYQOAMEESYVDLADKLKAGSTGMKVGFRA 416
DB 358 LSRTGIENLAKLEDRKEPWLVLVYAPWCPYQOAMEESYVDLADKLKAGS-GYKVGKFRADG 416
QY 417 EQKEFAKSELQGSFPTILFPFKHSSRPTIKYPSEKRDVDSLMAFVNALR 466
DB 417 DQKGFSGQELQGSFPTILFPFKHSSRP-IKYPSEKRDVDSLMAFVNALR 465

RESULT 3
Q672Q8_LYCSES PRELIMINARY; PRT; 461 AA.
AC Q672Q8;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Adenyllyl-sulfate reductase precursor.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
OX NCBI_TaxID=4081;
RN NUCLEOTIDE SEQUENCE.
RP Coker J.S., Vian A., Davies E.;
RA "Identification, accumulation, and functional prediction of novel
RT tomato transcripts systemically upregulated after fire damage.";
RL Physiol. Plantarum 124:311-322 (2005) .
DR EMBL; AY568717; AAU03359.1; -; mRNA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0016671; F:oxidoreductase activity, acting on sulfur g. . .; IEA.
DR GO; GO:0016740; P:transferrase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0019421; P:sulfate reduction, APS pathway; IEA.
DR InterPro; IPR004508; APS_reduc.
DR InterPro; IPR002500; PAPS_reduct.
DR InterPro; IPR006662; ThioRedox.
DR InterPro; IPR006663; ThioRedox_dom2.
DR InterPro; IPR012336; ThioRedoxin-like.
DR Pfam; PF01507; PAPS_reduct; 1.
DR PRINTS; PR00421; THIOREDOXIN.
DR TIGRFAMs; TIGR00424; APS_reduc; 1.
KW Transit peptide.
FT TRANSIT 1 71 Potential.
SQ SEQUENCE 461 AA; 51146 MW; F7D6BAD38914A799 CRC64;

```

Query Match 74.9%; Score 1832.5; DB 2; Length 461;

Best Local Similarity 75.1%; Pred. No. 2e-137;  
Matches 353; Conservative 49; Mismatches 50; Indels 19; Gaps 8;

QY 1 MALAFTSSIS-----SIGAPTSTFPSSPKLPQIGSIRISERPIGGAVNFNLSQRRSLVKPV 54  
DB 1 MALFTSSSIAHGLSSSSSY--EQPKVSLQTFQPLDRPQLSLTVLNRRAVAKPL 58  
QY 55 NABPPKDSIVPLAATTIVASASETKE-EDFEQIASLDNASPLEIMDRALDKFGNDIAI 113  
DB 59 YABPKKNDISVPSAATTIVAPEGVSEAEDEFKAKELQNASPLEVMDKALEKFGDDIAI 118  
QY 114 AFSGAEDVALIEYAKLTGPRFRVFSLDTGRLNPETVQLDPAVEKHGIRIEMFPDAVE 173  
DB 119 AFSGAEDVALIEYAHUTGPRFRVFSLDTGRLNPETVQLDPAVEKHGIRIEMFPDSEV 178  
QY 174 QALVRSGKLFSEYEDGHCRCRVRKVRPLRRALKGLRAWITGQRKQDSQPTSEIIPVQV 233  
DB 179 QALVRTGKLFSEYEDGHCRCRVRKVRPLRRALKGLRAWITGQRKQDSQPTSEIIPVQV 238  
QY 234 DPAFEGMDGIGSLVRKPNVANYKHGDINWFLRTMNPVNSLHAKGVSVIGCEPCTRPVL 293  
DB 239 DPFEGELGAGSLVRKPNVANYKHGDINWFLRTMNPVNSLHAKGVSVIGCEPCTRPVL 298  
QY 294 PGQHERGRRWEDAKAKECGLHKGNIWFLRTMNPVNSLHAKGVSVIGCEPCTRPVL 353  
DB 299 PGQHERGRRWEDAKAKECGLHKGNIWFLRTMNPVNSLHAKGVSVIGCEPCTRPVL 350  
QY 354 VNLSTRTGIENLAKLDRKPELVLVYAPWCPCQAMEESYVDLADKLAGSTGMKVGKFR 413  
DB 351 IVTLSPGVENLVKLEDRREPVLVYAPWCPCQAMEESYVDLADKLAGSTGMKVGKFR 409  
QY 414 ADGEKFAKSELQLSFPTILFFPKHSSRPPIKYPSEKRDVDSLMAFYNALR 466  
DB 410 ADGQKAFQAEELQLSFPTILFFPKHSSK-AIKYPSEKRDVDSLMAFYNALR 461

RESULT 4  
ID Q39619 CATRO PRELIMINARY; PRT; 464 AA.  
AC Q39619;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE PAPS-reductase-like protein precursor.  
GN Names:pat2neu;  
OS Catharanthus roseus (Rosy periwinkle) (Madagascar periwinkle).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
OC asterids; lamids; Gentianales; Apocynaceae; Rauvolfioideae; Vincetaceae;  
OC Catharanthus.  
OX NCBI\_TaxID=4058;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=9918320; PubMed=10082930; DOI=10.1016/S0167-4838(98)00266-0;  
RA Prior A., Uhrig J.F., Heins L., Wicemann A., Lillig C.H., Stoltze C.,  
RA Soll J., Schwenn J.D.;  
RT "Structural and kinetic properties of adenylyl sulfate reductase from  
Catharanthus roseus cell cultures";  
RL Biochim. Biophys. Acta 1430:25-38(1999).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RA Prior A., Schwenn J.D.;  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U63784; AAB05871.2; -, mRNA.  
DR F01; T10065; T10065.  
DR GO; GO:0005489; P:electron transporter activity; IEA.  
DR GO; GO:0016491; P:oxidoreductase activity; IEA.  
DR GO; GO:0016671; P:oxidoreductase activity, acting on sulfur g. . .; IEA.  
DR GO; GO:0016740; P:transferase activity; IEA.  
DR GO; GO:0008652; P:amino acid biosynthesis; IEA.  
DR GO; GO:0019344; P:cysteine biosynthesis; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR GO; GO:0008152; P:metabolism; IEA.

DR GO; GO:0019421; P:sulfate reduction, APS pathway; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR004508; APS\_reduc.  
DR InterPro; IPR002500; PAPS\_reduc.  
DR InterPro; IPR006662; ThioRed.  
DR InterPro; IPR006663; ThioRedox\_dom2.  
DR Pfam; PF01507; PAPS\_reduct; 1.  
DR Pfam; PF00085; ThioRedoxin; 1.  
DR PRINTS; PRO0421; THIOREDOXIN.  
DR TIGRFAMs; TIGR00424; APS\_reduc; 1.  
KW Transit peptide.  
FT TRANSIT 1 59 Potential.  
SQ SEQUENCE 464 AA; 51625 MW; A46FASAD551CB222 CRC64;

Query Match 74.8%; Score 1829; DB 2; Length 464;  
Best Local Similarity 75.1%; Pred. No. 3.9e-137;  
Matches 353; Conservative 48; Mismatches 59; Indels 10; Gaps 6;

QY 1 MALAFTSSIS---APTSTFPSSPKLPQIGSIRISERPIGGAVNFNLSQRRSLVKPVNAE 57  
DB 1 MALAFTSSIAHGLSSSSPEQTAAAAAQFGSFPQPLDRPHTISPSVNVSRRLAVKGINAE 60  
QY 58 PRKDSIVPLAATTIVASASETKE-EDFEQIASLDNASPLEIMDRALDKFGNDIAIAPS 116  
DB 61 PKRNEISVPSAATTIVAPEVEEKVDYDEYKJADELQNASPLEIMDKSLAKFGNDIAIAPS 120  
QY 117 GAEDVALIEYAKLTGPRFRVFSLDTGRLNPETVQLDPAVEKHGIRIEMFPDAVEQAL 176  
DB 121 GAEDVALIEYAHUTGPRFRVFSLDTGRLNPETVQLDPAVEKHGIRIEMFPDAVEQAL 180  
QY 177 VRSKGLFSYEDGHCRCRVRKVRPLRRALKGLRAWITGQRKQDSQPTSEIIPVQVQPA 236  
DB 181 VRSKGLFSYEDGHCRCRVRKVRPLRRALKGLRAWITGQRKQDSQPTSEIIPVQVQPV 240  
QY 237 FEGMDGGISLVKPNVANYKHGDINWFLRTMNPVNSLHAKGVSVIGCEPCTRPVLPGQ 296  
DB 241 FEGMDGGISLVKPNVANYKHGDINWFLRTMNPVNSLHAKGVSVIGCEPCTRPVLPGQ 300  
QY 297 HEREGRRWEDAKAKECGLHKGNIWFLRTMNPVNSLHAKGVSVIGCEPCTRPVLPGQ 356  
DB 301 HEREGRRWEDAKAKECGLHKGNIWFLRTMNPVNSLHAKGVSVIGCEPCTRPVLPGQ 356  
QY 357 LSRTGIENLAKLDRKPELVLVYAPWCPCQAMEESYVDLADKLAGSTGMKVGKFRADG 416  
DB 357 LSRTGIENLAKLDRKPELVLVYAPWCPCQAMEESYVDLADKLAGSTGMKVGKFRADG 415  
QY 417 EOKFAKSELQLSFPTILFFPKHSSRPPIKYPSEKRDVDSLMAFYNALR 466  
DB 416 DQKAFQAEELQLSFPTILFFPKHSSK-IPKYPSEKRDVDSLMAFYNALR 464

RESULT 5  
APRI ARATH STANDARD; PRT; 465 AA.  
ID APRI ARATH STANDARD; PRT; 465 AA.  
AC P92979; O48886; Q39248;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-MAY-2005 (Rel. 47, Last annotation update)  
DE 5'-adenylyl sulfate reductase 1, chloroplast precursor (EC 1.8.4.9)  
DE (Adenosine 5'-phosphosulfate 5'-adenylyl sulfate sulfoltransferase 1)  
DE (APS sulfoltransferase 1) (ThioRedoxin independent APS reductase 1)  
DE (3'-phosphoadenosine-5'-phosphosulfate reductase homolog 19) (PAPS  
DE reductase homolog 19) (Pth-19).  
GN Name=APR1; Synonyms=PRH19; OrderedLocusNames=At4g04610;  
GN ORFNames=F4H6.13;  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=cv. Columbia;







QY 297 HEREGRWEDAKAKCGLHKGKGVKQKEDVNGVNGLSQSHANGDATTVPDIENSPNVN 356  
 Db 298 HEREGRWEDAKAKCGLHKGKNIKE-----NTNGNATANVNGTA-TVADIENSENVN 350  
 QY 357 LSRGTGLENLAKLEDRKPEMLVVLVYAPWPCVQCOAMEESYVDLADKLKAGTGMVKYCKPRADG 416  
 Db 351 LSRQGIENLMLKLENKREANIVLVYAPWPCVQCOAMEASFDELADKLKGS-GVKAKPRADG 409  
 QY 417 EQKEFAKSEIQLGSPFTILFPFKHSSRTTIKYPSEKRDVDSLMAFVNALR 466  
 Db 410 DQXDFAKKEIQLGSPFTILFPFKHSSSRP-IKYPSEKRDVDSLFLNLR 458

RESULT 9  
 APR3 ARATH  
 ID \_APR3 ARATH STANDARD; PRT; 458 AA.  
 AC P92980; Q38948;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE 5'-adenylylsulfate reductase 3, chloroplast precursor (BC 1.8.4.9)  
 DE (adenosine 5'-phosphosulfate 3', chloroplast precursor)  
 DE (APS sulfotransferase 3) (Thioredoxin independent APS reductase 3)  
 DE (3'-phosphoadenosine-5'-phosphosulfate reductase homolog 26) (PAPS reductase homolog 26) (Prh-26).  
 GN Name=APR3; Synonyms=PRH26; OrderedLocusNames=At4g21990;  
 GN ORFNames=F1N20..90;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=97075173; PubMed=8917599; DOI=10.1073/pnas.93.23.13377;  
 RA Gutierrez-Marcos J.F., Roberts M.A., Campbell E.I., Wray J.L.;  
 RT "Three members of a novel small gene-family from Arabidopsis thaliana able to complement functionally an Escherichia coli mutant defective in PAPS reductase activity encode proteins with a thioredoxin-like domain and 'APS reductase' activity.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:13377-13382(1996).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=97075174; PubMed=8917600; DOI=10.1073/pnas.93.23.13383;  
 RA Setya A., Murillo M., Leustek T.;  
 RT "Sulfate reduction in higher plants: molecular evidence for a novel 5'-adenylylsulfate reductase.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:13383-13388(1996).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=cv. Columbia;  
 RA Chen Y.C., Leustek T.;  
 RT "Three genomic clones from Arabidopsis thaliana encoding 5'-adenylylsulfate reductase.";  
 RL (er) Plant Gene Register PGR98-030.  
 RN [4]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=20083488; PubMed=10617199; DOI=10.1038/47134;  
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G., Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terry N., Harris B., Ansong W., Brandt P., Grivell L.A., Rieger M., Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M., Kreis M., Delsen M., Fugdomenech P., Watson M., Schmidheini T., Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I., Vos P., Hohseil J., Zimmermann W., Wedler H., Ridley P., Langham S.-A., McCullagh B., Bilham L., Robben J., Van der Schueren J., Grynoprez B., Chuang Y.-J., Vandenbuesche F., Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E., Weitzenecker T., Bothe G., Ramsperger U., Hilbert H., Braun M., Holzer E., Brandt A., Peters S., van Staveren M., Dirkee W.,

RA Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P., Berniseier S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H., De Keyser A., Buyschaert C., Gielen J., Villarroel R., De Clercq R., Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S., Clark L., Doggett J., Hall S., Kay M., Lennard N., McLeay K., Mayes R., Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S., Borkova D., Blecker H., Scharfe M., Grimm M., Loehner T.-H., Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fartmann B., Grandrath K., Dauner D., Herzl A., Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E., Massenot O., Quigley F., Clabaud G., Muendlein A., Felber R., Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S., Chefdor F., Cooke R., Berger C., Monfort A., Casacuberta E., Gibbons T., Weber N., Vandenbol M., Bagues M., Terol J., Torres A., Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T., Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C., Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S., Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K., Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L., Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J., Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J., Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D., Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K., Kramer J., Fulton L., Mardis E., Dance M., Pepin K., Hillier L.W., Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D., Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C., Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C., Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R., Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S., Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A., Chen E., Marra M.A., Martienssen R., McCombie W.R.;  
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.";  
 RL Nature 402:769-777(1999).  
 RN [5]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=22954850; PubMed=14593172; DOI=10.1126/science.1088305;  
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J., Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F., Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G., Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J., Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y., Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P., Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F., Hayashizaki Y., Johnson-Hopson C., Huan V.W., Iida K., Karnes M., Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J., Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T., Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y., Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;  
 RT "Empirical analysis of transcriptional activity in the Arabidopsis genome.";  
 RL Science 302:842-846(2003).  
 CC -I- FUNCTION: Reduces sulfate for Cys biosynthesis. Substrate preference is adenosine-5'-phosphosulfate (APS) >> 3'-phosphoadenosine-5'-phosphosulfate (PAPS). Uses glutathione or DTT as source of protons.  
 CC -I- CATALYTIC ACTIVITY: AMP + sulfite + glutathione disulfide = adenylyl sulfate + 2 glutathione.  
 CC -I- ENZYME REGULATION: Stimulated by sodium sulfate > ammonium sulfate (by similarity).  
 CC -I- SUBCELLULAR LOCATION: Chloroplast (potential).  
 CC -I- TISSUE SPECIFICITY: Leaves, roots and stem.  
 CC -I- INDUCTION: By sulfate starvation.  
 CC -I- DOMAIN: The C-terminal domain may function as glutaredoxin and mediates the interaction of the enzyme with glutathione (GSH). Active in GSH-dependent reduction of hydroxyethylsulfide, cysteine, dehydroascorbate, insulin disulfides and ribonucleotide reductase (by similarity).  
 CC -I- SIMILARITY: Belongs to the APS reductase family.  
 CC -I- SIMILARITY: Contains 1 thioredoxin domain.  
 CC -I- CAUTION: Ref.2 sequence differs from that shown due to a frameshift in position 409.







Qy 231 VQVDPAFEGMDGGISLVKNNPVANVKGHDWNLFTWVNVVNSHLHAKGVYSGCEPCTR 290  
Dy 233 VQVDPVEFLDGGVGLVKNPLANVEGADVWNLFTWVNVNHLHAKGVYSGCEPCTR 292  
Qy 291 PVLPGQHERGRWWEADAKCEGLHKGNYKQKEEDVNGLSQSHANGDATVTPDIEN 350  
Dy 293 PVLPGQHERGRWWEADAKCEGLHKGNI---KEED-----GAADSKPAAVQSEIFE 341  
Qy 351 SPNVVNLSTRTIENLAKLEDRKEPWLVLVYAPWPCYQCAEESVYDLADKLKAGTGMKVG 410  
Dy 342 SNNVVALSRGCVENLLKLENKEAWLVLYAPWPCYQCAEESVYDLADKLKAGTGMKVG 400  
Qy 411 KFRADGQKPEAKSELQSGFPTLLFPKPKSSRTIKYKPEKRDVDSLMAPFNALR 466  
Dy 401 KFRADGQKPEAKSELQSGFPTLLFPKRAPR-AIKYPSEHRDVSILMSFVNLLR 455

## RESULT 11

APR2 ARATH STANDARD; PRT; 454 AA.  
AC P92981; O04215; O22554; Q38947;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-MAY-2005 (Rel. 47, Last annotation update)  
DE 5'-adenylsulfate reductase 2, chloroplast precursor (SC 1.8.4.9)  
DE (Adenosine 5'-phosphosulfate 5'-adenylsulfate sulfotransferase 2)  
DE (APS sulfotransferase 2) (Thioredoxin independent APS reductase 2)  
DE (3'-phosphoadenosine-5'-phosphosulfate reductase homolog 43) (PAPS  
DE reductase homolog 43) (Prh-43).  
GN Name=APR2; Synonym=APSR, PRH43; OrderedLocusNames=Atlg62180;  
GN ORFNames=F19K23.11;  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=97075173; PubMed=8917599; DOI=10.1073/pnas.93.23.13377;  
RA Gutierrez-Marcos J.F., Roberts M.A., Campbell E.I., Wray J.L.;  
RT "three members of a novel small gene-family from Arabidopsis thaliana  
RT able to complement functionally an Escherichia coli mutant defective  
RT in PAPS reductase activity encode proteins with a thioredoxin-like  
RT domain and 'APS reductase' activity.";  
RL Proc. Natl. Acad. Sci. U.S.A. 93:13377-13382 (1996).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=cv. Columbia; TISSUE=Seedling;  
RA Min B., Shin K.W., Ye X., Lee S.Y.;  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=cv. Columbia;  
RA Gutierrez-Marcos J.F., Campbell E.I., Wray J.L.;  
RT "A fourth member of the APS reductase gene family in Arabidopsis  
RT thaliana.";  
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=cv. Columbia;  
RA Chen Y.C., Leustek T.;  
RT "Three genomic clones from Arabidopsis thaliana encoding 5'-  
RT adenylylsulfate reductase.";  
RL (er) Plant Gene Register FGR98-030.  
RN [5]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=cv. Columbia;  
RX MEDLINE=21016719; PubMed=11130712; DOI=10.1038/35048500;  
RA Theologis A., Ecker J.R., Palm C.J., Federapfel N.A., Kaul S.,  
RA White O., Alonso J., Altieri H., Araujo R., Bowman C.I., Brooks S.Y.,  
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,

RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,  
RA Miliutescher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
RA Sakano H., Sallberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
RA Uterback T., Van Aken S., Vaysberg M., Vysocka V.S., Walker M.,  
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
RT thaliana.";  
RL Nature 408:816-820 (2000).  
RN [6]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
RC STRAIN=cv. Columbia;  
RX MEDLINE=22954850; PubMed=14593172; DOI=10.1126/science.1088305;  
RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,  
RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,  
RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,  
RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,  
RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,  
RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,  
RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,  
RA Havaehizaki Y., Johnson-Hopson C., Heuan V.W., Iida K., Karnes M.,  
RA Khan S., Koesena E., Ishida J., Jiang P.X., Jones T., Kawai J.,  
RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,  
RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,  
RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;  
RT "Empirical analysis of transcriptional activity in the Arabidopsis  
RT genome.";  
RL Science 302:842-846 (2003).  
RN [7]  
RP NUCLEOTIDE SEQUENCE OF 45-454.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=97075174; PubMed=8917600; DOI=10.1073/pnas.93.23.13383;  
RA Setya A., Murillo M., Leustek T.;  
RT "Sulfate reduction in higher plants: molecular evidence for a novel  
RT 5'-adenylsulfate reductase.";  
RL Proc. Natl. Acad. Sci. U.S.A. 93:13383-13388 (1996).  
CC -1- FUNCTION: Reduces sulfate for Cys biosynthesis. Substrate  
CC preference is adenosine-5'-phosphosulfate (APS) >> 3'-  
CC phosphoadenosine-5'-phosphosulfate (PAPS). Uses glutathione or DTT  
CC as source of protons.  
CC -1- CATALYTIC ACTIVITY: AMP + sulfite + glutathione disulfide =  
CC adenylyl sulfate + 2 glutathione.  
CC -1- ENZYME REGULATION: Stimulated by sodium sulfate > ammonium sulfate  
CC (by similarity).  
CC -1- SUBCELLULAR LOCATION: Chloroplast (Potential).  
CC -1- TISSUE SPECIFICITY: Leaves and stem.  
CC -1- INDUCTION: By sulfate starvation.  
CC -1- DOMAIN: The C-terminal domain may function as glutaredoxin and  
CC mediates the interaction of the enzyme with glutathione (GSH).  
CC Active in GSH-dependent reduction of hydroxyethylsulfide,  
CC cystine, dehydroascorbate, insulin disulfides and ribonucleotide  
CC reductase (by similarity).  
CC -1- SIMILARITY: Belongs to the APS reductase family.  
CC -1- SIMILARITY: Contains 1 thioredoxin domain.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
CC EMBL; US3866; AAC49563.1; -; mRNA.  
CC EMBL; AF023167; AAB80957.1; -; mRNA.  
CC EMBL; U96045; AAB57688.1; -; mRNA.  
CC EMBL; AF016283; AAC26980.1; -; Genomic DNA.  
CC EMBL; AC000375; AAB60764.1; -; Genomic DNA.  
CC EMBL; AF360192; AAK25902.1; -; mRNA.



Db	349	PAVSLTRTGIEENLLRLLENRAEPMLVLVYAPWCFCQMEASYVELAEKLAGS-GVKVAK	466
Qy	412	FRADGEQKEFAKSELQSGSPFTILFFPKPHSRPTIKYPSEKRDDVDLSLMAFVNALR	466
Db	408	FRADGEQKEFAQAELQSQSFPTVLLFPQRTARP-IKYPSKRDDVDLSLAFVNSLR	461
 RESULT 14 QSEUC8_MAIZE			
ID	QSEUC8_MAIZE PRELIMINARY;	PRT; 466 AA.	
AC	QSEUC8;		
DT	10-MAY-2005 (TrEMBLrel. 30, Created)		
DT	10-MAY-2005 (TrEMBLrel. 30, Last sequence update)		
DT	10-MAY-2005 (TrEMBLrel. 30, Last annotation update)		
DE	Adenosine 5'-phosphosulfate reductase 2.		
GN	Name=APRUZ;		
OS	Zea mays (Maize).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
OC	PACCAD clade; Panicoideae; Andropogoneae; Zea.		
OX	NCBI_taxID=4577;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RA	Jung R.;		
RL	Submitted (AUG-2004) to the EMBL/GenBank/DDBJ databases.		
DR	EMBL; AY739297; AAW63052.1; -; mRNA.		
DR	InterPro; IPR004508; APS_reduc.		
DR	InterPro; IPR002500; PAPS_reduc.		
DR	InterPro; IPR006662; Thiored.		
DR	InterPro; IPR006663; Thioredox_dom2.		
DR	Pfam; PF01507; PAPS_reduc; 1.		
DR	Pfam; PF00085; Thioredoxin; 1.		
DR	PRINTS; PR00421; THIOREDOXIN.		
DR	TIGRFAMS; TIGR00424; APS_reduc; 1.		
SQ	SEQUENCE 466 AA; 50404 MW; 84EE8D6435E9B238 CRC64;		
 Query Match 68.0%; Score 1663; DB 2; Length 466; Best Local Similarity 67.2%; Pred. No. 6.9e-124; Matches 322; Conservative 49; Mismatches 82; Indels 26; Gaps 5			
Qy	1	MALAFTSSISAPTSTPFSSEPFLPOLGSIKIRISERPIGGAVNFNLISORSLVKPVNAEPPR	60
Db	1	MASATTSISSHPAAL---RDVKAARIIGGLQQQASVAPSAARAGQRARVRSLRAAEPA	57
Qy	61	KDSIVPLAAITTVAS-----ASETKEEDFEQIASDLDNASPLETMDRALDKFGNDIAIAF	115
Db	:	: : : : : : : : : : : : : : : : :	
Db	58	RQPVAASAAAAAAPVADAEEAALAAVDYEALAELEGASPLETMDRALAMFGSEIATAF	117
Qy	116	SGAEDVALIYAKLTGRPFVFSLDTGRLPETTYQLFDNAVBEKHYGIRIENYFPDPDAVEVOA	175
Db	118	SGAEDVALIYAKLTGRPFVFSLDTGRLPETYQLFDKVBEKHYGIRIENYFPDASEVQE	177
Qy	176	LVSCKGLFSFYEDGHOBCCRVKVRPLRRALKGLRAMITGORKQDOSPTCRSEIPWQVDP	235
Db	178	LVRTKGLFSFYEDGHQBCCRVKVRPLRRALRGRLAMITGORKQDOSPTCRASIPLVQVDP	237
Qy	236	AFGMDGGIGSLVKWNPNVANVKGHDIIWNFLTMMNVPNVNSLHAKGYVISGCPCPTRPVLPG	295
Db	238	SPEGLDGAGSLVKWNPNVANVDGKDIIWTLFRTMDVPVNALHAQGYVISGCPCPTRPVLPG	297
Qy	296	OHEREGRWWEADAKACEGLHKGNVKOOKE-----EDVNGNGLSQSHANGDATTTVPD	347
Db	298	OHEREGRWWEADAKACEGLHKGNIDKQGAAAAAAAAPRTANGNG-----SAGAPD	348
Qy	348	IFNSPNVNLSRTGIENLALEDRKEPWLVLVYAPWCPCPYQAMEESYVDLDKLAGSGTM	407
Db	349	IPESPVVSLTRAGVENLLRLRESRAEPMLVLVYAPWCPCPYQAMEASYVELAEKLAGSGGV	408
Qy	408	KVGKFRADGQKEFAKSELQSGSPFTILFFPKPHSRPTIKYPSEKRDDVDLSLMAFVNALR	466
Db	409	KVARFRADGQKEFFAQAELOQSFPVTLLFPSTRARP-IKYPSKRDDVDLSLAFVNSLR	466



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2006, 14:37:29 ; Search time 49 Seconds  
(without alignments)  
786.263 Million cell updates/sec

Title: US-10-731-525-8  
Perfect score: 2445  
Sequence: 1 MALAFTSSISAPTSTFSSSE.....KYPSEKRDVDSLMFVNALR 466

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/6\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/H\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/PCUS\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/RE\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2445	100.0	466	2	US-09-720-318A-8
2	1923	78.7	472	2	US-09-720-318A-6
3	1811.5	74.1	465	2	US-09-720-318A-12
4	1788.5	73.1	463	2	US-09-720-318A-11
5	1645.5	67.3	463	2	US-09-720-318A-10
6	1295.5	53.0	299	2	US-09-720-318A-4
7	1262.5	51.6	233	2	US-09-720-318A-2
8	811.5	33.2	260	2	US-09-328-352-5189
9	745.5	30.5	269	2	US-09-252-991A-30515
10	540	22.1	240	2	US-09-902-540-15169
11	407.5	16.7	251	2	US-09-540-236-2957
12	347	14.2	236	2	US-10-126-279-28
13	347	14.2	236	2	US-10-286-606-28
14	322	13.2	247	2	US-10-126-279-29
15	322	13.2	247	2	US-10-286-606-29
16	313.5	12.8	254	2	US-10-126-279-30
17	313.5	12.8	254	2	US-10-286-606-30
18	307.5	12.6	248	2	US-09-134-001C-4299
19	305.5	12.5	254	2	US-10-126-279-27
20	305.5	12.5	254	2	US-10-286-606-27
21	256.5	10.5	261	2	US-09-962-357-8
22	250.5	10.2	155	2	US-09-710-279-308
23	206.5	8.4	248	2	US-09-543-681A-7091
24	202	8.3	260	2	US-09-489-039A-9955
25	154	6.3	638	1	US-08-557-122A-38
26	154	6.3	638	2	US-09-262-666-38
27	154	6.3	645	2	US-09-538-092-920

28	150	6.1	496	2	US-09-807-258-27	Sequence 27, Appl
29	145	5.9	511	1	US-08-557-122A-4	Sequence 4, Appli
30	145	5.9	511	2	US-09-262-666-4	Sequence 4, Appli
31	145	5.9	515	1	US-08-557-122A-3	Sequence 3, Appli
32	145	5.9	515	1	US-08-557-122A-34	Sequence 34, Appli
33	145	5.9	515	2	US-09-262-666-3	Sequence 3, Appli
34	145	5.9	515	2	US-09-262-666-34	Sequence 34, Appli
35	145	5.9	3052	1	US-08-557-122A-26	Sequence 26, Appl
36	145	5.9	3052	2	US-09-262-666-26	Sequence 26, Appl
37	140	5.7	489	2	US-09-807-258-10	Sequence 10, Appl
38	139.5	5.7	366	2	US-08-984-919A-11	Sequence 11, Appl
39	139.5	5.7	486	2	US-08-984-919A-33	Sequence 33, Appl
40	139.5	5.7	470	2	US-08-984-919A-55	Sequence 55, Appl
41	139.5	5.7	495	2	US-08-984-919A-47	Sequence 47, Appl
42	139.5	5.7	522	2	US-09-368-588-2	Sequence 2, Appli
43	139	5.7	747	2	US-09-999-833A-459	Sequence 459, App
44	139	5.7	747	2	US-10-020-445A-459	Sequence 459, App
45	138.5	5.7	368	2	US-08-781-420-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1  
US-09-720-318A-8  
; Sequence 8, Application US/09720318A  
; Patent No. 6730827  
; GENERAL INFORMATION:  
; APPLICANT: Falco, Saverio Carl  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167-C  
; CURRENT APPLICATION NUMBER: US/09/720,318A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 1998-07-14  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 8  
; LENGTH: 466  
; TYPE: PRT  
; ORGANISM: Glycine max  
; US-09-720-318A-8

Query Match	100.0.0%;	Score	2445;	DB 2;	Length	466;			
Best Local Similarity	100.0.0%;	Pred. No.	3.4e-260;						
Matches	466;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	MALAFTSSISAPTSTPPSSEPKLPQIGSRISRPIGGAVNLSQRRLSKPVNAEPPR	60						
Db	1	MALAFTSSISAPTSTPPSSEPKLPQIGSRISRPIGGAVNLSQRRLSKPVNAEPPR	60						
Qy	61	KDSIVPLAATTIVASASETKEDFEQIASDLNAPLEIMDRALDKFGNDIAFSGAED	120						
Db	61	KDSIVPLAATTIVASASETKEDFEQIASDLNAPLEIMDRALDKFGNDIAFSGAED	120						
Qy	121	VALIEYAKLTGRPFVFSLDTGRLNPETQYLPDAVEKHGIRIETMPPDAVEQVALVRSK	180						
Db	121	VALIEYAKLTGRPFVFSLDTGRLNPETQYLPDAVEKHGIRIETMPPDAVEQVALVRSK	180						
Qy	181	GLFSFYEDGHQECRCVRKVRPLRALKGLRAMITGQRKDQSPCTRSEIPVVQVDPAFEGM	240						
Db	181	GLFSFYEDGHQECRCVRKVRPLRALKGLRAMITGQRKDQSPCTRSEIPVVQVDPAFEGM	240						
Qy	241	DGGISGLVKWNPVANYKGHDIWNFLRTMNPVNSLHAKGVYSTGCEPCTRPVLPQGHORE	300						
Db	241	DGGISGLVKWNPVANYKGHDIWNFLRTMNPVNSLHAKGVYSTGCEPCTRPVLPQGHORE	300						
Qy	301	GRWWEDAKAKECGLHKGNNVKQKQEDVNGSLGSHANGDATTVDPDIFNSPNVNLNLSRT	360						
Db	301	GRWWEDAKAKECGLHKGNNVKQKQEDVNGSLGSHANGDATTVDPDIFNSPNVNLNLSRT	360						
Qy	361	GIENLAKLEDKRPWLVLVYAPWCPCYQAMEESYVDLADKLKAGSTGMKVGKFRADGEQKE	420						







; Sequence 2, Application US/09720318A  
; Patent No. 6730827  
; GENERAL INFORMATION:  
; APPLICANT: Falco, Saverio Carl  
; APPLICANT: Allen, Stephen M.  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167-C  
; CURRENT APPLICATION NUMBER: US/09/720,318A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 1998-07-14  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 2  
; LENGTH: 293  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (91)  
; OTHER INFORMATION: Xaa = any amino acid  
US-09-720-318A-2

Query Match 51.6%; Score 1262.5; DB 2; Length 293;  
Best Local Similarity 77.3%; Pred. No. 3.4e-130; Mismatches 27; Indels 15; Gaps 4;  
Matches 235; Conservative 27;

QY 167 FPDAVEQALVRSKGLSFYEDGHQECRCVRKVRPLRRALKGLRAWITGQRKQSPGTRS 226  
DB 1 FPDASEQVELVIRKGLSFYEDGHQECRCVRKVRPLRRALKGLRAWITGQRKQSPGTRA 60

QY 227 EIPVQVDPAFEGVDGIGISLVKWNPNVANKGHDINFLRTMNPVNSLHAKGVYSGICE 286  
DB 61 SIPVQVDPSEFGLDGGAGSLVKNPNVANKGDIWFLRTMDVPVNTLHAQGVYSGICE 120

QY 287 PCTRPVLPGOHEREGRWWEDAKACEGLHKNVKKQKE----EDVNGGLSQSHANGDA 342  
DB 121 PCTRPVLPGOHEREGRWWEDAKACEGLHKNIDKDAQAAAPRSANGNG-----S 171

QY 343 TTVDPIDFNSPNVNLRTGIENLAKLEDKRPMLVLVYAPWCPYQOAMEESYVDLADKLA 402  
DB 172 AGAPDIFESAVVSLTGTGIENLLRLNRAEPMLVLVYAPWCPYQOAMEASYVELAEKLA 231

QY 403 GSTGMKVGKFRADGEQKEFAKSELQLGSFPTILFPFKHSSRPRTIKYPSEKRDVSLMAFV 462  
DB 232 GS-GVKVAKFRADGEQKFAELQLOSFPTVLLFPGRTRP-IKYPSEKRDVSLLAFF 289

QY 463 NALR 466  
DB 290 NSLR 293

RESULT 8  
US-09-328-352-5189  
; Sequence 5189, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 5189  
; LENGTH: 260  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-5189

Query Match 33.2%; Score 811.5; DB 2; Length 260;  
Best Local Similarity 57.3%; Pred. No. 1.4e-80; Mismatches 62; Indels 9; Gaps 3;  
Matches 150; Conservative 41;

QY 61 KOSIVPLAATTIVASASETKEDFEIOIASDLDNASPLEIMDRALDKFGNDIAIFSGAED 120  
DB 8 RPSLLPGTPTWTVPTI-----DIVDALAAEYAAKSPREILELALSQOG-EIALSFGAED 61

QY 121 VALIEYAKLTGRPRVFSLDTGRNLNPETYQLFDVAKEXHYGIRIETMFPDPAVEQALVRSK 180  
DB 62 VVLIDNASRLGKPRVFSLDTGRLHPETQFIETVRKHYNINIEICFPDAEAVQSMVNEK 121

QY 181 GLPSFYEDGHQECRCVRKVRPLRRALKGLRAWITGQRKQSPGTRSIEIPVQVDPAFEGM 240  
DB 122 GLPSFFKDGHOECGIRKVKQPLRKLATLDGWITGQRKQSPGTRTEIPVQADAGFSGP 181

QY 241 DGGIGISLVKWNPNVANKGHDINFLRTMNPVNSLHAKGVYSGICEPCTRPVLPQOHERE 300  
DB 182 G---KQLIKYNPLANWSSADVSYIRNMEIPYNPLHERGFVSGCEPCTRPVLPNQHERE 238

QY 301 GRWWEDAKACEGLHKNVKKQ 322  
DB 239 GRWWEEATQKEGCLHAGNLKK 260

RESULT 9  
US-09-252-991A-30515  
; Sequence 30515, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 30515  
; LENGTH: 269  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-30515

Query Match 30.5%; Score 745.5; DB 2; Length 269;  
Best Local Similarity 54.6%; Pred. No. 2.9e-73; Mismatches 130; Conservative 46; Mismatches 59; Indels 3; Gaps 1;  
Matches 130;

QY 83 DFEQIASDLDNASPLEIMDRALDKFGNDIAIFSGAEDVALIEYAKLTGRPRVFSLDTG 142  
DB 31 DLPALASSLADKSPQDILKAAFBHFGDELWISFGAEDVVLVDMAWLNKRVKVFSLDTG 90

QY 143 RLNPETYQLFDVAKEXHYGIRIETMFPDPAVEQALVRSKGLFSFYEDGHQECRCVRKVRPL 202  
DB 91 RLHPETVRFIDQVREHYGIAIDVLSPDRLLEPLVKEKGLFSFYRDGHCCEGCKIRKIEPL 150

QY 203 RRALKGLRAWITGQRKQSPGTRSIEIPVQVDPAFEGMDGIGISLVKWNPNVANKGHDIV 262  
DB 151 KRKLAVRAWATGQRDQSPGTRSQAVALTEIDGAFSTPE---KPLYKFNPLSSMTSEEW 207

QY 263 NFLRTMNPVNSLHAKGVYSGICEPCTRPVLPQOHEREGRWWEDAKACEGLHKNV 320  
DB 208 GYINMLPELNPVNSLHERGYISIGCEPCTRPVLPNQHEREGRWWEDAKACEGLHAGNL 265

RESULT 10  
US-09-902-540-15169  
; Sequence 15169, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.

; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 15169  
; LENGTH: 240  
; TYPE: PRT  
; ORGANISM: Myxococcus xanthus  
US-09-902-540-15169

Query Match 22.1%; Score 540; DB 2; Length 240;  
Best Local Similarity 43.5%; Pred. No. 1.1e-50;  
Matches 114; Conservative 43; Mismatches 75; Indels 30; Gaps 8;  
QY 58 PPKDSIVPLAATTIVASATKEEDFEQIASLDNASPLEIMDRALDKFGNDIAIAFS- 116  
DB 3 PPQPQ-----ASALLAAEAVKDAP-----AQDL-----LAWTER---RFGARAAIASSF 44  
QY 117 GAEDVALIEYAKLTGRPFVFSLDTGRNLNPTVQLFDVAVEKHGIRIEYMFDPDAVEQAL 176  
DB 45 GVDMVLIDLARQAPSLRFLTLDTGRLPPEYELMEVVRKRYGVTVETTFPERARVEAL 104  
QY 177 VRSGKLFSPYE--DGHQCCRVKVRPLRLKALGLRAWITGQRKQDQSPGTRSEIPVQVD 234  
DB 105 ESTNGYFSPQSLKARACCAIRKVEPLSRALAGQQAQWVTLGRREQSV-TETDVAITLVD 163  
QY 235 PAPEGMDGGTSLVKNNPVANVGHDIWNFLRTMNPVNSLHAKGYVSGICEPCTRPVLP 294  
DB 164 SAH-----GLKLNPLATWSRDIWAYVRAKSVYPYNALHGRGYPISGCAPCTRAVKP 215  
QY 295 GQHERGREGWWEDAKAKECGLH 316  
DB 216 YEDERAGRWWSAENRECCGLH 237

RESULT 11  
US-09-540-236-2957  
; Sequence 2957, Application US/09540236  
; Patent No. 6673910  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR  
; FILE REFERENCE: 2709.2005-001  
; CURRENT APPLICATION NUMBER: US/09/540,236  
; CURRENT FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 3840  
; SEQ ID NO 2957  
; LENGTH: 251  
; TYPE: PRT  
; ORGANISM: M.catarhalis  
US-09-540-236-2957

Query Match 16.7%; Score 407.5; DB 2; Length 251;  
Best Local Similarity 40.9%; Pred. No. 4.6e-36;  
Matches 88; Conservative 33; Mismatches 77; Indels 17; Gaps 5;  
QY 106 KFGNDIAIAFSGAEDVALIEYAKLTGRPFVFSLDTGRNLNPTVQLFDVAVEKHGIRIEYMFDPDAVEQAL 164  
DB 51 KPASSLQ-----AEDVITDALANSHLSTEIFILQIGRLNAETLKLDIVKACLYLINFK 105  
QY 165 YMPFPAVEQALVRSKGLFSFYEDG--HQCCRVKVRPLRLKALGLRAWITGQRKQDQSP 222  
DB 106 TYEPHPHADVAVYANHGLNAFYSGDLRKLCCFIRKVEPLNRLALVDADAWLTQORREQSV 165  
QY 223 GTSEIPVQVDPAFEGMDGGTSLVKNNPVANVGHDIWNFLRTMNPVNSLHAKGYVS 282  
DB 166 -TRTELNLAEITDLA-----RGIKYNPFDQETDVMWAYTLTKNIPNELYHOGYPS 216

QY 283 IGCEPCTRPVLPQOHEREGRWWWEDAKAKECGLH 317  
DB 217 IGCEPCTMPVKQGEDIRAGRWWHWHKONKECGLH 251

RESULT 12  
US-10-126-279-28  
; Sequence 28, Application US/10126279  
; Patent No. 6858213  
; GENERAL INFORMATION:  
; APPLICANT: Bertozzi, Carolyn  
; APPLICANT: Williams, Spencer J.  
; APPLICANT: Mougous, Joseph  
; TITLE OF INVENTION: Mycobacterial Sulfation Pathway Proteins  
; FILE REFERENCE: BERK-012  
; CURRENT APPLICATION NUMBER: US/10/126,279  
; CURRENT FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/285,394  
; PRIOR FILING DATE: 2001-04-20  
; PRIOR APPLICATION NUMBER: 60/345,953  
; PRIOR FILING DATE: 2001-10-26  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 28  
; LENGTH: 236  
; TYPE: PRT  
; ORGANISM: Mycobacterium smegmatis  
US-10-126-279-28

Query Match 14.2%; Score 347; DB 2; Length 236;  
Best Local Similarity 37.1%; Pred. No. 1.9e-29;  
Matches 92; Conservative 34; Mismatches 106; Indels 16; Gaps 7;  
QY 70 TTIVASATKEEDFEQIASLDNASPLEIMDRALDKFGNDIAIAFSGAEDVALIEYAKL 129  
DB 2 TDVTTSTENELRELAERGAELADASAELLRTDHFHGGNYVVA-SNMQDVLVEMAAK 60  
QY 130 TGRPFVFSLDTGRNLNPTVQLFDVAVEKHGIRIEYMFDP-AVEQALVRSKGLFSFYED 188  
DB 61 VRPGVDVLFDITGYHFAETIGTRDAVEYVDVHVNVVTVERTVAEQDELIGKILFA-RDP 119  
QY 189 GHQCCRVKVRPLRLKALGLRAWITGQRKQDQSPGTRSEIPVQVDPAFEGMDGGTSLV 248  
DB 120 G--ECCRLKRWPLTNALKYSAWVTGIRREAP-TRANAPLISWDNAF-----GLV 168  
QY 249 KNPVANVGHDIWNFLRTMNPVNSLHAKGYVSGICEPCTRPVLPQOHEREGRWWWEDA 308  
DB 169 KINPIAAWTDQMNYIDANGILVNLPLVYEGYPSIGCAPCTSKPIPCADPRSGR--WAGL 226  
QY 309 KAKECGLH 316  
DB 227 SKTECGLH 234

RESULT 13  
US-10-286-606-28  
; Sequence 28, Application US/10286606  
; Patent No. 6863895  
; GENERAL INFORMATION:  
; APPLICANT: Bertozzi, Carolyn  
; APPLICANT: Williams, Spencer J.  
; APPLICANT: Mougous, Joseph  
; TITLE OF INVENTION: Mycobacterial Sulfation Pathway Proteins  
; FILE REFERENCE: BERK-012CIP  
; CURRENT APPLICATION NUMBER: US/10/286,606  
; CURRENT FILING DATE: 2002-10-31  
; PRIOR APPLICATION NUMBER: 60/285,394  
; PRIOR FILING DATE: 2001-04-20  
; PRIOR APPLICATION NUMBER: 60/345,953  
; PRIOR FILING DATE: 2001-10-26  
; PRIOR APPLICATION NUMBER: 10/126,279

```
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 236
; TYPE: PR1
; ORGANISM: Mycobacterium smegmatis
US-10-286-606-28

Query Match      14.2%; Score 347; DB 2; Length 236;
Best Local Similarity 37.1%; Pred. No. 1.9e-29;
Matches 92; Conservative 34; Mismatches 106; Indels 16; Gaps 7;

Qy 70 TTIVASASEYKEEDFEQIADSLDNASPLEIMDRALDKFGNDIAIFSGAEDVALIEYAKL 129
Db 2 TDTVTSTENELRELAERGAELADASABELLRTWDEHFGGNYVVA-SNMQDAVLVEMAAK 60

Qy 130 TGRPFVFSLDTGRNLNPEYQLFDAVEKHGIRIEMFPD-AVEVQALVRSKGLFSFYED 188
Db 61 VRPGVDVFLDTGTHFAETIGTRDAVEAVDVHVVTPTVAEQDELLGKNLFA-RDP 119

Qy 189 GHQCCRVKVRPLRRALKGLRAWITGQRKQSPGTRSEIPVVQVDPAPFGMDGGIGSLV 248
Db 120 G--ECCRLRKVPPLTNALKGYSAMVTGIRRVKAP--TRANAPLISWDNAF-----GLV 168

Qy 249 KWNPVANVKHDIWNFLRTMNVFVNSLHAKGYVSGICEPCTRPVLPQOHEREGRWWEDEA 308
Db 169 KINPIAAWTEDMQNYIDANGILYNPLVYEGYSGICAPCTSKPIPGADPRSGR--WAGL 226

Qy 309 KAKECGLH 316
Db 227 SKTECGLH 234

RESULT 14
US-10-126-279-29
; Sequence 29, Application US/10126279
; Patent No. 6858213
; GENERAL INFORMATION:
; APPLICANT: Bertozzi, Carolyn
; APPLICANT: Williams, Spencer J.
; TITLE OF INVENTION: Mycobacterial Sulfation Pathway Proteins
; FILE OF INVENTION: and Methods of Use Thereof
; FILE REFERENCE: BERK-012
; CURRENT APPLICATION NUMBER: US/10/126,279
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/285,394
; PRIOR FILING DATE: 2001-04-20
; PRIOR FILING DATE: 2001-10-26
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 247
; TYPE: PR1
; ORGANISM: Mycobacterium avium
US-10-126-279-29

Query Match      13.2%; Score 322; DB 2; Length 247;
Best Local Similarity 35.8%; Pred. No. 1.2e-26;
Matches 86; Conservative 36; Mismatches 92; Indels 26; Gaps 8;

Qy 88 ASDLDNASPLEIMDRALDKFGN-----DIAIAFSGAEDVALIEYAKLTGRPPRVF 137
Db 21 AAELGASASDVLRWTDTEFGVNGPRGMATCNVVA-SSMQEAVLIDLAALKVRPGVPV 79

Qy 138 SLDTGRNLNPEYQLFDAVEKHGIRIEMFPD-AVEVQALVRSKGLFSFYEDGHQECRV 196
Db 80 FLDTGYHFAETIGTRDAIESYDILRVNLVTPHESVAEQDKLLGKDLFA-RDPG--ECRL 136

Qy 197 RKVRLRRALKGLRAWITGQRKQSPGTRSEIPVVQVDPAPFGMDGGIGSLVKNPVANV 256
Db 137 RKVAPLGLKTLRGYSAMVTGLRRSEA-ATRAPAPVIGFDEGFK-----LVKVPNPMATW 187

Qy 257 KGHDIWNFLRTMNVFVNSLHAKGYVSGICEPCTRPVLPQOHEREGRWWEDEAKAKECGLH 316
Db 188 TDEDVQNYIDEHNVNPLVYEGYSGICAPCTAKPLAGADPRSGR--WQGLAKTECGLH 245

Search completed: February 18, 2006, 14:38:50
Job time : 50 secs
```

```
Db 137 RKVAPLGLKTLRGYSAMVTGLRRSEA-ATRAPAPVIGFDEGFK-----LVKVPNPMATW 187
Qy 257 KGHDIWNFLRTMNVFVNSLHAKGYVSGICEPCTRPVLPQOHEREGRWWEDEAKAKECGLH 316
Db 188 TDEDVQNYIDEHNVNPLVYEGYSGICAPCTAKPLAGADPRSGR--WQGLAKTECGLH 245

RESULT 15
US-10-286-606-29
; Sequence 29, Application US/10286606
; Patent No. 6863895
; GENERAL INFORMATION:
; APPLICANT: Bertozzi, Carolyn
; APPLICANT: Williams, Spencer J.
; TITLE OF INVENTION: Mycobacterial Sulfation Pathway Proteins
; FILE OF INVENTION: and Methods of Use Thereof
; FILE REFERENCE: BERK-012CIP
; CURRENT APPLICATION NUMBER: US/10/286,606
; PRIOR FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: 60/285,394
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/345,953
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: 10/126,279
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 247
; TYPE: PR1
; ORGANISM: Mycobacterium avium
US-10-286-606-29

Query Match      13.2%; Score 322; DB 2; Length 247;
Best Local Similarity 35.8%; Pred. No. 1.2e-26;
Matches 86; Conservative 36; Mismatches 92; Indels 26; Gaps 8;

Qy 88 ASDLDNASPLEIMDRALDKFGN-----DIAIAFSGAEDVALIEYAKLTGRPPRVF 137
Db 21 AAELGASASDVLRWTDTEFGVNGPRGMATCNVVA-SSMQEAVLIDLAALKVRPGVPV 79

Qy 138 SLDTGRNLNPEYQLFDAVEKHGIRIEMFPD-AVEVQALVRSKGLFSFYEDGHQECRV 196
Db 80 FLDTGYHFAETIGTRDAIESYDILRVNLVTPHESVAEQDKLLGKDLFA-RDPG--ECRL 136

Qy 197 RKVRLRRALKGLRAWITGQRKQSPGTRSEIPVVQVDPAPFGMDGGIGSLVKNPVANV 256
Db 137 RKVAPLGLKTLRGYSAMVTGLRRSEA-ATRAPAPVIGFDEGFK-----LVKVPNPMATW 187

Qy 257 KGHDIWNFLRTMNVFVNSLHAKGYVSGICEPCTRPVLPQOHEREGRWWEDEAKAKECGLH 316
Db 188 TDEDVQNYIDEHNVNPLVYEGYSGICAPCTAKPLAGADPRSGR--WQGLAKTECGLH 245

Search completed: February 18, 2006, 14:38:50
Job time : 50 secs
```

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2006, 14:48:35 ; Search time 168 Seconds  
(without alignments)  
1158.979 Million cell updates/sec

Title: US-10-731-525-8

Perfect score: 2445

Sequence: 1 MALAFTSSISAPTSTPSSSE.....KYPSEKRDVDSLMAFVNALR 466

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA Main.\*

1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep.\*

2: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep.\*

3: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep.\*

4: /cgn2\_6/ptodata/1/pubppaa/US10A\_PUBCOMB.pep.\*

5: /cgn2\_6/ptodata/1/pubppaa/US10B\_PUBCOMB.pep.\*

6: /cgn2\_6/ptodata/1/pubppaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2445	100.0	466	4	US-10-731-525-8
2	2445	100.0	466	4	US-10-762-049-8
3	1951	79.8	470	4	US-10-424-599-277584
4	1951	79.8	475	4	US-10-425-114-43990
5	1923	78.7	472	4	US-10-731-525-6
6	1923	78.7	472	4	US-10-762-049-6
7	1811.5	74.1	465	4	US-10-731-525-12
8	1811.5	74.1	465	4	US-10-762-049-12
9	1788.5	73.1	463	4	US-10-731-525-11
10	1788.5	73.1	463	4	US-10-762-049-11
11	1645.5	67.3	463	4	US-10-731-525-10
12	1645.5	67.3	463	4	US-10-762-049-10
13	1639.5	67.1	475	4	US-10-437-363-126234
14	1622.5	66.4	473	5	US-10-739-930-8328
15	1597.5	65.3	459	4	US-10-169-667A-2
16	1509.5	61.7	386	4	US-10-425-115-267885
17	1444.5	59.1	407	4	US-10-425-115-267882
18	1422.5	58.2	339	4	US-10-435-115-267886
19	1295.5	53.0	299	4	US-10-731-525-4
20	1295.5	53.0	299	4	US-10-762-049-4
21	1262.5	51.6	293	4	US-10-731-525-2
22	1262.5	51.6	293	4	US-10-762-049-2
23	1241.5	50.8	323	4	US-10-425-114-56536
24	1110.5	45.4	227	4	US-10-424-599-278242
25	1045.5	42.8	251	4	US-10-425-115-267884
26	999.5	40.9	252	4	US-10-425-115-267887
27	914	37.4	182	4	US-10-425-115-331871

28	871.5	35.6	213	4	US-10-767-701-42399	Sequence 42399, A
29	754	30.8	147	4	US-10-424-599-225761	Sequence 225761, A
30	731.5	29.9	235	4	US-10-369-493-13795	Sequence 13795, A
31	606	24.8	248	4	US-10-425-115-197638	Sequence 197638, A
32	539	22.0	230	4	US-10-369-493-19207	Sequence 19207, A
33	537	22.0	109	4	US-10-767-701-49099	Sequence 49099, A
34	454.5	18.6	184	4	US-10-424-599-212812	Sequence 212812, A
35	449.5	18.4	225	4	US-10-369-493-19679	Sequence 19679, A
36	410	16.8	158	4	US-10-425-115-340785	Sequence 340785, A
37	393.5	16.1	219	4	US-10-369-493-9785	Sequence 9785, A
38	370.5	15.2	220	4	US-10-369-493-11025	Sequence 11025, A
39	347	14.2	236	4	US-10-126-279-28	Sequence 28, Appl
40	347	14.2	236	4	US-10-286-606-28	Sequence 28, Appl
41	347	14.2	236	5	US-10-891-383-28	Sequence 28, Appl
42	325	13.3	196	4	US-10-369-493-17686	Sequence 17686, A
43	322	13.2	247	4	US-10-126-279-29	Sequence 29, Appl
44	322	13.2	247	4	US-10-286-606-29	Sequence 29, Appl
45	322	13.2	247	5	US-10-891-383-29	Sequence 29, Appl

## ALIGNMENTS

### RESULT 1

US-10-731-525-8

; Sequence 8, Application US/10731525

; Publication No. US20040121440A1

; GENERAL INFORMATION:

; APPLICANT: Falco, Saverio Carl

; APPLICANT: Allen, Stephen M.

; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins

; FILE REFERENCE: BB-1167-C

; CURRENT APPLICATION NUMBER: US/10/731,525

; CURRENT FILING DATE: 2003-12-09

; PRIOR APPLICATION NUMBER: US/09/720,318A

; PRIOR FILING DATE: 2000-12-21

; PRIOR APPLICATION NUMBER: 60/092,833

; PRIOR FILING DATE: 1998-07-14

; NUMBER OF SEQ IDS: 12

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 8

; LENGTH: 466

; TYPE: PRT

; ORGANISM: Glycine max

US-10-731-525-8

Query Match		100.0%;	Score 2445;	DB 4;	Length 466;
Best Local Similarity		100.0%;	Pred. No. 1.1e-222;		
Matches 466;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MALAFTSSISAPTSTPSSSEPKLPQIGSIRISERPTGGAVNFNLSQRSLVKPVNASPPR	60		
Db	1	MALAFTSSISAPTSTPSSSEPKLPQIGSIRISERPTGGAVNFNLSQRSLVKPVNASPPR	60		
Qy	61	KDSIVPLAATTIVASASSETKEEDFEQIASDLNASPLEIMDRALDKFGNDIAFAFSAED	120		
Db	61	KDSIVPLAATTIVASASSETKEEDFEQIASDLNASPLEIMDRALDKFGNDIAFAFSAED	120		
Qy	121	VALIEYAKLTGRPFVFLSDTGRNLNPNETYLQFAVEKHGYIRIEMPPDPAVEQALVRSK	180		
Db	121	VALIEYAKLTGRPFVFLSDTGRNLNPNETYLQFAVEKHGYIRIEMPPDPAVEQALVRSK	180		
Qy	181	GLFSFYEDGHQECRCVKRPLRRALKGLRAWITGQRKQSGPTRSRIPVVQVDPAFEGM	240		
Db	181	GLFSFYEDGHQECRCVKRPLRRALKGLRAWITGQRKQSGPTRSRIPVVQVDPAFEGM	240		
Qy	241	DGGIGSLVKWNPVANKGHDDIWNFLRTMNPVNSLHAKGVYSIGCEPCTRPVLPQGHRE	300		
Db	241	DGGIGSLVKWNPVANKGHDDIWNFLRTMNPVNSLHAKGVYSIGCEPCTRPVLPQGHRE	300		
Qy	301	GRWWEDAKAKECGLHKGNVKQKEEDVNGLSQSHANGDATTVDPIDFNSPNVNLNLSRT	360		
Db	301	GRWWEDAKAKECGLHKGNVKQKEEDVNGLSQSHANGDATTVDPIDFNSPNVNLNLSRT	360		

Qy 361 GIENLAKLEDKKEPWLVLVYAPWCPYCOAMEESYVDLADKLKAGSTGMKVGKFRADGEOKE 420  
Db 361 GIENLAKLEDKKEPWLVLVYAPWCPYCOAMEESYVDLADKLKAGSTGMKVGKFRADGEOKE 420  
Qy 421 FAKSELQSGSPFTTLFFPKHSSRPTIKYPSKRDVDSLMAFVNALR 466  
Db 421 FAKSELQSGSPFTTLFFPKHSSRPTIKYPSKRDVDSLMAFVNALR 466

RESULT 2  
US-10-762-049-8  
; Sequence 8, Application US/10762049  
; Publication No. US20040139492A1  
; GENERAL INFORMATION:  
; APPLICANT: Falco, Saverio Carl  
; APPLICANT: Allen, Stephen M.  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167-C  
; CURRENT APPLICATION NUMBER: US/10/762,049  
; CURRENT FILING DATE: 2004-01-21  
; PRIOR APPLICATION NUMBER: US/09/720,318A  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 1998-07-14  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 8  
; LENGTH: 466  
; TYPE: PRT  
; ORGANISM: Glycine max  
US-10-762-049-8

Query Match 100.0%; Score 2445; DB 4; Length 466;  
Best Local Similarity 100.0%; Pred. No. 1.1e-222;  
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALAFTSSISAPTSTFFSSSEPKLPQIGSIRISERPIGGAVNFNLSQRRSLVKPNABPPR 60  
Db 1 MALAFTSSISAPTSTFFSSSEPKLPQIGSIRISERPIGGAVNFNLSQRRSLVKPNABPPR 60

Qy 61 KDSIVPLAATTIVASASTKEEDFEQIASDLNADNASPLEIMDRALDKFGNDIAIAFSAED 120  
Db 61 KDSIVPLAATTIVASASTKEEDFEQIASDLNADNASPLEIMDRALDKFGNDIAIAFSAED 120

Qy 121 VALIEYAKLTGRPRFRVSLDTGRNLNPTYQLFDAVEKHGIRIEMPPDAVEQALVRSK 180  
Db 121 VALIEYAKLTGRPRFRVSLDTGRNLNPTYQLFDAVEKHGIRIEMPPDAVEQALVRSK 180

Qy 181 GLFSFYEDGHOECRCRVRKRLRALKGLRAWITGQRKQSPGTRSEIPVQVDPAPFGM 240  
Db 181 GLFSFYEDGHOECRCRVRKRLRALKGLRAWITGQRKQSPGTRSEIPVQVDPAPFGM 240

Qy 241 DGGIGSLVKMNPVANVKGHD1WNFLRTMNVPVNSLHAKGVYSIGCEPCTRVLPGQHERE 300  
Db 241 DGGIGSLVKMNPVANVKGHD1WNFLRTMNVPVNSLHAKGVYSIGCEPCTRVLPGQHERE 300

Qy 301 GRWWEDAKACEGLHKGNVKQKEEDVNGNLSQSHANGDATTVPDIFNSPNVNLNLSRT 360  
Db 301 GRWWEDAKACEGLHKGNVKQKEEDVNGNLSQSHANGDATTVPDIFNSPNVNLNLSRT 360

Qy 361 GIENLAKLEDKKEPWLVLVYAPWCPYCOAMEESYVDLADKLKAGSTGMKVGKFRADGEOKE 420  
Db 361 GIENLAKLEDKKEPWLVLVYAPWCPYCOAMEESYVDLADKLKAGSTGMKVGKFRADGEOKE 420

Qy 421 FAKSELQSGSPFTTLFFPKHSSRPTIKYPSKRDVDSLMAFVNALR 466  
Db 421 FAKSELQSGSPFTTLFFPKHSSRPTIKYPSKRDVDSLMAFVNALR 466

RESULT 3  
US-10-424-599-277584  
; Sequence 277584, Application US/10424599

; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 277584  
; LENGTH: 470  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_92682C.1.pep  
US-10-424-599-277584

Query Match 79.8%; Score 1951; DB 4; Length 470;  
Best Local Similarity 80.3%; Pred. No. 8.3e-176;  
Matches 382; Conservative 38; Mismatches 40; Indels 16; Gaps 7;

Qy 1 MALAF-----TSSISAPTSTF-----PSSEPKLPQIGSIRISERPIGGAVNFNLSQRRSL 50  
Db 1 MALAVTTSSSSAAAASSFFSRLLGSSSDAKAPQIGSFPPERSLVSSVVVNTQRRSL 60

Qy 51 VKPVNAEPKRDSTIVPLAATTIVASASTKEEDFEQIASDLNADNASPLEIMDRALDKFGND 110  
Db 61 VRPLNAEPQRNDSIVPLAA--TIVAPEVEKEEDFEQIAKOLENASPLEIMDRALDKFGND 119

Qy 111 IATAFSAEDVALIEYAKLTGRPRFRVSLDTGRNLNPTYQLFDAVEKHGIRIEMPPDA 170  
Db 120 IATAFSAEDVALIEYAKLTGRPRFRVSLDTGRNLNPTYQLFDAVEKHGIRIEMPPDA 179

Qy 171 VEQVALVRSKGLFSFYEDGHOECRCRVRKRLRALKGLRAWITGQRKQSPGTRSEIPV 230  
Db 180 VEQVALVRSKGLFSFYEDGHOECRCRVRKRLRALKGLRAWITGQRKQSPGTRSEIPV 239

Qy 231 VQVDPAPFEGDGGIGSLVKMNPVANVKGHD1WNFLRTMNVPVNSLHAKGVYSIGCEPCTR 290  
Db 240 VQVDPVFEGLDGGIGSLVKMNPVANVNGLD1WSFLRTMDVPVNSLHSGYVYSIGCEPCTR 299

Qy 291 PVLPGQHEREGRWWEDAKACEGLHKGNVKQKEEDVNGNLSQSHANGDATTVPDIFN 350  
Db 300 PVLPGQHEREGRWWEDAKACEGLHKGNTKHEDAAQLNGNGASQ--ANGSA--TVADIFN 356

Qy 351 SPNVNLSRTGIENLAKLEDKKEPWLVLVYAPWCPYCOAMEESYVDLADKLKAGSTGMKVG 410  
Db 357 SQDVVLSRSGIENLAKLENKKEPWLVLVYAPWCRFCQAMEESYVDLAEKLAGS--GVKVA 415

Qy 411 KFRADGQKEFAKSELQSGSPFTTLFFPKHSSRPTIKYPSKRDVDSLMAFVNALR 466  
Db 416 KFRADGQKEYAKTELQSGSPFTTLFFPKHSSQF--INKYPSKRDVDSLTAFAVNALR 470

RESULT 4  
US-10-425-114-43990  
; Sequence 43990, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28

[illegible]

```
Db 182 VEQALVTRTGLSFYEDGHECCRVKVRPLRRALKGLKAWITGQRKQSPGTRSEIPI 241
Qy 231 VQVDPAPFEGMDGGIGSLVKNPVAVKGDHINWFLRTMNPVNSLHAKGYVSGCBPCTR 290
Db 242 VQVDPVPEGLDGGIGSLVKNPVAVNNGLDIWNFLRTMNPVNSLHAKGYVSGCBPCTR 301
Qy 291 PVLPGQHERGRWWEDAKAKCEGLHKGNGVKQKQKEEDVNGGLSQSHANGDATTVPDIEN 350
Db 302 PVLPGQHERGRWWEDAKAKCEGLHKGNGLKQEDAAQLNGNGTSG--GNLSA-TVADIFI 358
Qy 351 SPNVNLSRGTGIENLAKLEDRKEPWLVLVYAPWCPCYQCAAMEESYVDLADKLKAGSTGMKVG 410
Db 359 SQNVVLSRSGIENLAKLENKEHMLVVLVYAPWCPCYQCAAMEESYVDLAEKLARS-GVKVA 417
Qy 411 KFRADGQKFAKSELQSGFPTILFPFKSSRRPTIKYPSEKRDVDSLMAFVNALR 466
Db 418 KFRADGQKFAKSELQSGFPTILFPFKSSRRPTIKYPSEKRDVDSLMAFVNALR 472

RESULT 7
US-10-731-525-12
; Sequence 12, Application US/10731525
; Publication No. US20040121440A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/10731525
; CURRENT FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: US/09/720,318A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-731-525-12

Query Match 74.1%; Score 1811.5; DB 4; Length 465;
Best Local Similarity 73.9%; Pred. No. 1.4e-162;
Matches 351; Conservative 48; Mismatches 57; Indels 19; Gaps 7;

Qy 1 MALAFTSSISAPTSTFPSS-----EPKLPQIGSIRISERPTIGGAVNFNLS-QRRSLVKPV 54
Db 1 MAMSVNVSSSSSGIINSRFGVSLPEKVSQIGSLRLDLDRVHVAVPSLNLGKSSSVKPL 60
Qy 55 NAEPPRKDSIVPLAATTIVASSETKE---EDFQIASDLNAPLEIMDRALDKFGNDI 111
Db 61 NAEPPRKDSIVPLAATTIVASSETKE---EDFQIASDLNAPLEIMDRALDKFGNDI 120
Qy 112 AIAFSGAEDVALIEYAKLTGRPRFVSLDTGRNLMPETYQLFDAVEKHGIRIETMFPDVA 171
Db 121 AIAFSGAEDVALIEYAKLTGRPRFVSLDTGRNLMPETYQLFDAVEKHGIRIETMFPDVA 180
Qy 172 EVQALVRSKGLFSFYEDGHECCRVKVRPLRRALKGLKAWITGQRKQSPGTRSEIPV 231
Db 181 EVQALVRSKGLFSFYEDGHECCRVKVRPLRRALKGLKAWITGQRKQSPGTRSEIPV 240
Qy 232 QVDPAPFEGMDGGIGSLVKNPVAVNNGLDIWNFLRTMNPVNSLHAKGYVSGCBPCTR 291
Db 241 QVDPAPFEGMDGGIGSLVKNPVAVNNGLDIWNFLRTMNPVNSLHAKGYVSGCBPCTR 300
Qy 292 VLPQHERGRWWEDAKAKCEGLHKGNGVKQKQKEEDVNGGLSQSHANGDATTVPDIEN 351
Db 301 VLPQHERGRWWEDAKAKCEGLHKGNGVKQKQKEEDVNGGLSQSHANGDATTVPDIEN 358
Qy 351 SPNVNLSRGTGIENLAKLEDRKEPWLVLVYAPWCPCYQCAAMEESYVDLADKLKAGSTGMKVG 410
Db 359 SQNVVLSRSGIENLAKLENKEHMLVVLVYAPWCPCYQCAAMEESYVDLAEKLARS-GVKVA 417
Qy 411 KFRADGQKFAKSELQSGFPTILFPFKSSRRPTIKYPSEKRDVDSLMAFVNALR 466
Db 418 KFRADGQKFAKSELQSGFPTILFPFKSSRRPTIKYPSEKRDVDSLMAFVNALR 472
```

```
Query Match 74.1%; Score 1811.5; DB 4; Length 465;
Best Local Similarity 73.9%; Pred. No. 1.4e-162;
Matches 351; Conservative 48; Mismatches 57; Indels 19; Gaps 7;

Qy 1 MALAFTSSISAPTSTFPSS-----EPKLPQIGSIRISERPTIGGAVNFNLS-QRRSLVKPV 54
Db 1 MAMSVNVSSSSSGIINSRFGVSLPEKVSQIGSLRLDLDRVHVAVPSLNLGKSSSVKPL 60
Qy 55 NAEPPRKDSIVPLAATTIVASSETKE---EDFQIASDLNAPLEIMDRALDKFGNDI 111
Db 61 NAEPPRKDSIVPLAATTIVASSETKE---EDFQIASDLNAPLEIMDRALDKFGNDI 120
Qy 112 AIAFSGAEDVALIEYAKLTGRPRFVSLDTGRNLMPETYQLFDAVEKHGIRIETMFPDVA 171
Db 121 AIAFSGAEDVALIEYAKLTGRPRFVSLDTGRNLMPETYQLFDAVEKHGIRIETMFPDVA 180
Qy 172 EVQALVRSKGLFSFYEDGHECCRVKVRPLRRALKGLKAWITGQRKQSPGTRSEIPV 231
Db 181 EVQALVRSKGLFSFYEDGHECCRVKVRPLRRALKGLKAWITGQRKQSPGTRSEIPV 240
Qy 232 QVDPAPFEGMDGGIGSLVKNPVAVNNGLDIWNFLRTMNPVNSLHAKGYVSGCBPCTR 291
Db 241 QVDPAPFEGMDGGIGSLVKNPVAVNNGLDIWNFLRTMNPVNSLHAKGYVSGCBPCTR 300
Qy 292 VLPQHERGRWWEDAKAKCEGLHKGNGVKQKQKEEDVNGGLSQSHANGDATTVPDIEN 351
Db 301 VLPQHERGRWWEDAKAKCEGLHKGNGVKQKQKEEDVNGGLSQSHANGDATTVPDIEN 358
Qy 351 SPNVNLSRGTGIENLAKLEDRKEPWLVLVYAPWCPCYQCAAMEESYVDLADKLKAGSTGMKVG 410
Db 359 SQNVVLSRSGIENLAKLENKEHMLVVLVYAPWCPCYQCAAMEESYVDLAEKLARS-GVKVA 417
```

```
Db 353 ENLVTLTSSRQGIENLAKLENKEPWIIVLYAPWCPCYQCAAMEASYDELADKLKAGS-GIKVAK 411
Qy 412 FRADGQKFAKSELQSGFPTILFPFKSSRRPTIKYPSEKRDVDSLMAFVNALR 466
Db 412 FRADGQKFAKSELQSGFPTILFPFKSSRRPTIKYPSEKRDVDSLMAFVNALR 465

RESULT 8
US-10-762-049-12
; Sequence 12, Application US/10762049
; Publication No. US20040139492A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/10762,049
; CURRENT FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: US/09/720,318A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-762-049-12

Query Match 74.1%; Score 1811.5; DB 4; Length 465;
Best Local Similarity 73.9%; Pred. No. 1.4e-162;
Matches 351; Conservative 48; Mismatches 57; Indels 19; Gaps 7;

Qy 1 MALAFTSSISAPTSTFPSS-----EPKLPQIGSIRISERPTIGGAVNFNLS-QRRSLVKPV 54
Db 1 MAMSVNVSSSSSGIINSRFGVSLPEKVSQIGSLRLDLDRVHVAVPSLNLGKSSSVKPL 60
Qy 55 NAEPPRKDSIVPLAATTIVASSETKE---EDFQIASDLNAPLEIMDRALDKFGNDI 111
Db 61 NAEPPRKDSIVPLAATTIVASSETKE---EDFQIASDLNAPLEIMDRALDKFGNDI 120
Qy 112 AIAFSGAEDVALIEYAKLTGRPRFVSLDTGRNLMPETYQLFDAVEKHGIRIETMFPDVA 171
Db 121 AIAFSGAEDVALIEYAKLTGRPRFVSLDTGRNLMPETYQLFDAVEKHGIRIETMFPDVA 180
Qy 172 EVQALVRSKGLFSFYEDGHECCRVKVRPLRRALKGLKAWITGQRKQSPGTRSEIPV 231
Db 181 EVQALVRSKGLFSFYEDGHECCRVKVRPLRRALKGLKAWITGQRKQSPGTRSEIPV 240
Qy 232 QVDPAPFEGMDGGIGSLVKNPVAVNNGLDIWNFLRTMNPVNSLHAKGYVSGCBPCTR 291
Db 241 QVDPAPFEGMDGGIGSLVKNPVAVNNGLDIWNFLRTMNPVNSLHAKGYVSGCBPCTR 300
Qy 292 VLPQHERGRWWEDAKAKCEGLHKGNGVKQKQKEEDVNGGLSQSHANGDATTVPDIEN 351
Db 301 VLPQHERGRWWEDAKAKCEGLHKGNGVKQKQKEEDVNGGLSQSHANGDATTVPDIEN 358
Qy 351 SPNVNLSRGTGIENLAKLEDRKEPWLVLVYAPWCPCYQCAAMEESYVDLADKLKAGSTGMKVG 410
Db 359 SQNVVLSRSGIENLAKLENKEPWIIVLYAPWCPCYQCAAMEASYDELADKLKAGS-GIKVAK 411
Qy 412 FRADGQKFAKSELQSGFPTILFPFKSSRRPTIKYPSEKRDVDSLMAFVNALR 466
Db 412 FRADGQKFAKSELQSGFPTILFPFKSSRRPTIKYPSEKRDVDSLMAFVNALR 465
```

```
RESULT 9
US-10-731-525-11
; Sequence 11, Application US/10731525
; Publication No. US20040121440A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
```



```
/ APPLICANT: Allen, Stephen M.
/ TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
/ FILE REFERENCE: BB-1167-C
/ CURRENT APPLICATION NUMBER: US/10/731,525
/ CURRENT FILING DATE: 2003-12-09
/ PRIOR APPLICATION NUMBER: US/09/720,318A
/ PRIOR FILING DATE: 2000-12-21
/ PRIOR APPLICATION NUMBER: 60/092,833
/ PRIOR FILING DATE: 1998-07-14
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: Microsoft Office 97
/ SEQ ID NO 11
/ LENGTH: 463
/ TYPE: PRT
/ ORGANISM: Catharanthus roseus
US-10-731-525-11

Query Match      73.1%; Score 1788.5; DB 4; Length 463;
Best Local Similarity 73.4%; Pred. No. 2.1e-160;
Matches 348; Conservative 49; Mismatches 58; Indels 19; Gaps 6;

QY 1 MALAFTSSIS--APTSTFPSSRPKLPOIGSIRISRPIGGAVNFNLSQRRSLVKPVNAE 57
Db 1 MALAFTSSTAIHGSLSSSFQTKAAAQFGSPQLDRPHTTISPVNRRRLAVKPINAE 60
QY 58 PPRKDSIVPLAATTIVASASETKE-EDFEQIASDLONASPLEIMDRALDKFGNDIAIAFS 116
Db 61 PKRNESIVPSAATTVAPEVEEKVDVEYKLADELQNASPLEIMDKSLAKFGNDIAIAFS 120
QY 117 GAEDVALIEYAKLTGRPRVFSLDTGRLNPETTYKFFDTVEKQYGIHIEYMFDPDAVEQAL 176
Db 121 GAEDVALIEYAKLTGRPRVFSLDTGRLNPETTYKFFDTVEKQYGIHIEYMFDPDAVEQAL 180
QY 177 VRSKGLFSPFYEDGHQCCRVKVRPLRRALKGLRAWITGORKDOQSPQTRSEIPVQVDPV 236
Db 181 VRSKGLFSPFYEDGHQCCRVKVRPLRRALKGLRAWITGORKDOQSPQTRSEIPVQVDPV 240
QY 237 FEGMDGGIGSLVKNPVPVANYKGHDINWFLRTMNPVNSLHAKGVYSGICEPTRPVLPQG 296
Db 241 FEGMDGGVSLVKNPVPVANYKGHDINWFLRAMDVPVNTLHSGQVYSGICEPTRPVLPQG 300
QY 297 HEREGRWWEDEKAKCEGLHKGNNVQKQK----BEDVNGLSQSHANGDATTVPDIPNSP 352
Db 301 HEREGRWWEDEKAKCEGLHKGDIKEGTLIWDGAVNGNG-----SDTIADIPDTN 351
QY 353 NVNLSRTGTENLAKLEDREKPEWLVLYAPWCPYCOAMEESYVDLADKLAGSTGMKVGKF 412
Db 352 NVTSLSRPGIENLLKLEEREAWLVDYAPWCRFCQAMEGSYLELAEKLAGS-GVRKVGKF 410
QY 413 RADGEQKEFAKSELQSGSPFTILFFPKHSSRPRTIKYPSEKRDVDSLMFAFNALR 466
Db 411 KADGDQKAPAQEQLQNLSSPTILFFPKHSSKP-IKYPSEKRDVDSLMFAFNALR 463

RESULT 10
US-10-762-049-11
/ Sequence 11, Application US/10762049
/ Publication No. US20040139492A1
/ GENERAL INFORMATION:
/ APPLICANT: Falco, Saverio Carl
/ APPLICANT: Allen, Stephen M.
/ TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
/ FILE REFERENCE: BB-1167-C
/ CURRENT APPLICATION NUMBER: US/10/762,049
/ CURRENT FILING DATE: 2004-01-21
/ PRIOR APPLICATION NUMBER: US/09/720,318A
/ PRIOR FILING DATE: 2000-12-21
/ PRIOR APPLICATION NUMBER: 60/092,833
/ PRIOR FILING DATE: 1998-07-14
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: Microsoft Office 97
/ SEQ ID NO 11
/ LENGTH: 463
/ TYPE: PRT
/ ORGANISM: Catharanthus roseus
US-10-731-525-11

Query Match      67.3%; Score 1645.5; DB 4; Length 463;
Best Local Similarity 68.8%; Pred. No. 7.9e-147;
Matches 327; Conservative 49; Mismatches 78; Indels 21; Gaps 8;

QY 1 MALAFTSSISAPTSTFPSSRPKLPOIGSIR--ISERPIGGAVNFNLSQRRSLVKPVNAEP 58
Db 1 MASA-TASIS--SHSIALRDLKAARIGAVRQVQVAPGLPATAPKQQRARVPLICNAE 57
QY 59 PRKDSIVPLAATTIVA-----SASETKEEDFEQIASDLONASPLEIMDRALDKFGNDIAIA 114
```

```

Db      58  PARPVSAASAASPVAPEEASAAVADYEAALQELVGASPLEIMDRALDMFGSEIAIA 117
Qy      115  FSGAEDVALIEYAKLTGRPFVFSLDTGRNLNPETYQLFDVAKHYGIRIEYMPDPDAVEVQ 174
Db      118  FSGAEDVALIEYAKLTGRPFVFSLDTGRNLNPETYQLFDVAKHYGIRIEYMPDPDAVEVQ 177
Qy      175  ALVRSKGLFSFYEDHQCCECRVRKVRPLRALKGLRAWITGORKDQSPGTRSRIPVQVD 234
Db      178  DLVRSKGLFSFYEDHQCCECRVRKVRPLRALKGLRAWITGORKDQSPGTRSRIPVQVD 237
Qy      235  PAFEGMDGGISLVKNNPVANVKGHDINWFLRTMNVVNSLHAKGYYSIGCEPCTRPVLP 294
Db      238  PSFEGLDGGAGSLIKWNPVANVDKDIWTLFRTMDVPNTLHAQGYYSIGCEPCTRPVLP 297
Qy      295  GQHERGRWWMEDAKAECGLHKGNVK---OQKEEDVNGNGLSQSHANGDATTVPDIFNS 351
Db      298  GQHERGRWWMEDATAKECGLHNGNIDKEGQAPKVGNGNGSABEASA-----PDIFQS 350
Qy      352  PNVNLSRTGIENLAKLEDRKEPWLVLVYAPWPCYCOAMEESYVDLADKLAGSTGMKVKG 411
Db      351  QAIVNLTRPGIENLLRENRAEPWLTVLYAPWPCYCOAMEASYVELAEKLSGS-GIKVAK 409
Qy      412  FRADGQEKFAKSELQSGPPTILFFPKHSSRPTIKYPSEKRDVDSLMAFVNALR 466
Db      410  FRADGQKPPAQAEQLQSPPTILLFPGRVTKEP-IKYPSEKRDVQSLAFVNSLR 463

```

## RESULT 12

```

US-10-762-049-10
; Sequence 10, Application US/10762049
; Publication No. US20040139492A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Stephen M.
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; CURRENT APPLICATION NUMBER: US/10762,049
; PRIOR FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: US/09/720,318A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-762-049-10

```

```

Query Match      67.3%; Score 1645.5; DB 4; Length 463;
Best Local Similarity 68.8%; Pred. No. 7.9e-147;
Matches 327; Conservative 49; Mismatches 78; Indels 21; Gaps 8;

Qy      1  MALAFTSSISAPTSTFPSSSEPKLQIGSIR--ISERPIGGAVNPNLSORRSLVKPVNAEP 58
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||

Db      1  MASA-TASIS--SHSIALDLKAARIGAVQQAAPVAVPAGLPATAPKQORAVRPLCAAE 57
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||

Qy      59  PRKDSIVPLAATTIVA---SASETKEEDPFOIASDLDNASPLEIMDRALDKFGNDIAIA 114
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||

Db      58  PARPVSAASAASPVAPEEASAAVADYEAALQELVGASPLEIMDRALDMFGSEIAIA 117
Qy      115  FSGAEDVALIEYAKLTGRPFVFSLDTGRNLNPETYQLFDVAKHYGIRIEYMPDPDAVEVQ 174
Db      118  FSGAEDVALIEYAKLTGRPFVFSLDTGRNLNPETYQLFDVAKHYGIRIEYMPDPDAVEVQ 177
Qy      175  ALVRSKGLFSFYEDHQCCECRVRKVRPLRALKGLRAWITGORKDQSPGTRSRIPVQVD 234
Db      178  DLVRSKGLFSFYEDHQCCECRVRKVRPLRALKGLRAWITGORKDQSPGTRSRIPVQVD 237
Qy      235  PAFEGMDGGISLVKNNPVANVKGHDINWFLRTMNVVNSLHAKGYYSIGCEPCTRPVLP 294

```

```

Db      238  PSFEGLDGGAGSLIKWNPVANVDKDIWTLFRTMDVPNTLHAQGYYSIGCEPCTRPVLP 297
Qy      295  GQHERGRWWMEDAKAECGLHKGNVK---OQKEEDVNGNGLSQSHANGDATTVPDIFNS 351
Db      298  GQHERGRWWMEDATAKECGLHNGNIDKEGQAPKVGNGNGSABEASA-----PDIFQS 350
Qy      352  PNVNLSRTGIENLAKLEDRKEPWLVLVYAPWPCYCOAMEESYVDLADKLAGSTGMKVKG 411
Db      351  QAIVNLTRPGIENLLRENRAEPWLTVLYAPWPCYCOAMEASYVELAEKLSGS-GIKVAK 409
Qy      412  FRADGQEKFAKSELQSGPPTILFFPKHSSRPTIKYPSEKRDVDSLMAFVNALR 466
Db      410  FRADGQKPPAQAEQLQSPPTILLFPGRVTKEP-IKYPSEKRDVQSLAFVNSLR 463

RESULT 13
US-10-437-963-126234
; Sequence 126234, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 126234
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_28800C.1.pep
US-10-437-963-126234

```

```

Query Match      67.1%; Score 1639.5; DB 4; Length 475;
Best Local Similarity 66.6%; Pred. No. 3e-146;
Matches 325; Conservative 53; Mismatches 75; Indels 35; Gaps 11;

Qy      1  MALAFTSSISAPTSTFPSSSEPKLQIGSIR--ISERPIGGAVNPNLSORRSLVKPVNAEP 58
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||

Db      1  MASA-TASISHSVAL--RDLKAARIGAVKQQAAPAAAGTAAAR-AQARAVRPLRAAE 56
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||

Qy      59  PRKDSIVPLAATTIVASASETKEE-----DFOIASDLDNASPLEIMDRAL 104
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||

Db      57  PARQ---PVSASAAAAAPAAVDAADAAAVDAPAPAVDYEAALQELVGASPLEIMDRAL 113
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||

Qy      105  DKFGNDIAIAPSGAEDVALIEYAKLTGRPFVFSLDTGRNLNPETYQLFDVAKHYGIRIE 164
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||

Db      114  AMFSDIAIAPSGAEDVALIEYAKLTGRPFVFSLDTGRNLNPETYQLFDVAKHYGIRIE 173
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||

Qy      165  YMFDPDAVEQVALVRSKGLFSFYEDHQCCECRVRKVRPLRALKGLRAWITGORKDQSPGT 224
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||

Db      174  YMFDPDAVEQVALVRSKGLFSFYEDHQCCECRVRKVRPLRALKGLRAWITGORKDQSPGT 233
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||

Qy      225  RSEIPVQVDPAPEGMDDGISLVKNNPVANVKGHDINWFLRTMNVVNSLHAKGYYSIG 284
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||

Db      234  RAALPVVQVDSFSGLAGGAGSLVKNNPVANVDGKVMVTFLRAMDVPVNLHAQGYYSIG 293
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||

Qy      285  CEPCTRPVLPQCHEREGRWWMEDAKAECGLHKGNVKQOKEEDVNGNGLSQSH-----ANG 340
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||

Db      294  CEPCTRPVLPQCHEREGRWWMEDAKAECGLHKGNI-----DDQGAAGAAAAHAKGANG 348
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||

Qy      341  DATT-VPDFNSPNVNLRSRTGIENLAKLEDRKEPWLVLVYAPWPCYCOAMEESYVDLAD 399
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||

Db      349  NGSAGAPDIFSSGVSILTRAGVENLLRLESRAEPWLVLVYAPWPCYCOAMEASYLELAE 408
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||
||| :||| : : :||| : : :||| : : :||| : : :||| : : :||| : : :|||

```



**THIS PAGE LEFT BLANK**

Result No.	Query %			DB	ID	Description
	Score	Match	Length			
1	250.5	10.2	155	6	US-10-793-626-308	Sequence 308, App
2	154	6.3	645	6	US-10-821-234-1409	Sequence 1409, Ap
3	139	5.7	747	6	US-10-131-826A-426	Sequence 426, App
4	127.5	5.2	505	6	US-10-821-234-1287	Sequence 1287, Ap
5	117	4.8	440	6	US-10-821-234-1282	Sequence 1282, Ap
6	110.5	4.5	324	6	US-10-878-556A-114	Sequence 114, App
7	107.5	4.4	269	7	US-11-000-463-344	Sequence 344, App
8	105.5	4.3	627	6	US-10-467-657-5432	Sequence 5432, Ap
9	104	4.3	176	6	US-10-821-234-860	Sequence 860, App
10	101.5	4.2	414	6	US-10-878-556A-1	Sequence 1, Appli
11	99	4.0	875	7	US-11-024-959-352	Sequence 352, App
12	91.5	3.7	183	6	US-10-467-962B-6	Sequence 6, Appli
13	91.5	3.7	844	6	US-10-763-712A-48	Sequence 48, Appl
14	90	3.7	310	7	US-11-156-084-227	Sequence 227, App
15	89.5	3.7	760	6	US-10-858-730-76	Sequence 76, Appl
16	88.5	3.6	732	6	US-10-954-468-51	Sequence 51, Appl
17	87.5	3.6	740	6	US-10-821-234-1464	Sequence 1464, Ap
18	85.5	3.5	408	6	US-10-763-712A-67	Sequence 67, Appl
19	85.5	3.5	456	6	US-10-763-712A-66	Sequence 66, Appl
20	85.5	3.5	610	7	US-11-069-643-22	Sequence 22, Appl
21	85	3.5	280	6	US-10-967-457-75	Sequence 75, Appl
22	85	3.5	440	6	US-10-525-710-52	Sequence 52, Appl
23	85	3.5	1102	7	US-11-098-686-10951	Sequence 10951, A
24	84.5	3.5	1735	7	US-11-040-472-13	Sequence 13, Appl
25	84	3.4	419	7	US-11-156-084-122	Sequence 122, App

```
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821.234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1409
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1409
```

```
Query Match      6.3%; Score 154; DB 6; Length 645;
Best Local Similarity 19.6%; Pred. No. 9.9e-06;
Matches 114; Conservative 82; Mismatches 192; Indels 194; Gaps 28;

QY   7  SSISAPSTPSSEPKLPQIGSIRISRPICGAVNFNLSQRRSLVKPVNAEPKQSIVP 66
Db   118 AKIDATASVSLASRFDSVGYPTTKLKK--GQAVDVEGSRTOBEIVAKVREVSOQPDWTPP 175
QY   67 LAATTIVASASETKEEDFEQIASDLD-----NASPLEIMDRALDKFGND 110
Db   176 PEVLVLT-----KENFDEVNDADILLVEFVAPWCCHCKLAPEYKAAKELSKRSPP 229
QY   111 IAIA-FSGAEDVALIEYAKLTGR-PRVFLSDTGRNLNPETYQLFDVAVEKHGIRIEYMF- 167
Db   230 IPLAKVDATAETDLAKRFDVSGYPTLKIFR--KGR--PYDY---NGPREKYGI-VDYMI 281
QY   168 ---PDAVEQALVRSGKLFSEYEDCHQECRCVRKPLRALKGLRAWITQKQDSFGT 224
Db   282 QSGPPSKSEILT---KQVQEFLLKDGDD-----ViliGVFKGES--- 316
QY   225 RSEIPVVQVDPAPFGMDGGI-----GSLV-----KW 250
Db   317 -----DPAYQOQDAANLREDYKPHHTFSTETAKFLKVSQGLVMQPEKFSKY 367
QY   251 NP---VANVRGHD-----IMNFLTMMNPV-----NSLHAKGYVSIIGCEPCTRPVLPQHE 298
Db   368 EPRSHMDVQSGTQDSAIKDFVLKVALPLVCHRKVSNDAKRYT-----RRPLVVVYYS 420
QY   299 -----REGRWV-----ED--AKACEGLHKGNVKQKKEEDV 328
Db   421 VDFSFYRAATQFWRSKVLEAKDFPEYTFATIADEEDYAGEVKDLGL-----SESGEDV 474
QY   329 N-----GNGLSQSHANGDATT-----VPIFNSPNVNLSRT 360
Db   475 NAILDESCKKFAPEPEFSDTLREFVTAFKGLKLPVKSQVPKNNKGPVKVVVGKT 534
QY   361 GIENLAKLEDRKEPWLVLVYAPWCPYCOAMEESYVDLADKLAGSTGMKVGKFRADGQKE 420
Db   535 -FDSI--VMDPKDVLIEFYAPWCGHCKQLEPVVNSLAKYKQKGLVIAKM--DATAND 589
QY   421 FAKSELQIGSPPTTLFPFKSSRRTIKYVPSEKRDVDSLMARV 462
Db   590 VPSDRYKVEGPTTYFAPSGDKKNPVKPEGGDRDLHLKFI 631
```

## RESULT 3

```
US-10-131-826A-426
; Sequence 426, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerriteen, Mary E.
; APPLICANT: Goddard, Audrey
```

```
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131.826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 426
; LENGTH: 747
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-131-826A-426
```

```
Query Match      5.7%; Score 139; DB 6; Length 747;
Best Local Similarity 21.8%; Pred. No. 0.00027;
Matches 105; Conservative 70; Mismatches 214; Indels 92; Gaps 23;

QY   27 GSIRISRPICGAVNFNLSQRRSLVKPVNAEP-----RKDSIVPLAA 69
Db   177 GLLRI-----GAVNCGDDRMCLCMKGVNSVPSLIFPRSGMAPVKYHGDRLSKESLVSPAM 230
QY   70 TTIVASASETKEEDFEQIASDLDNASPLEIMDRALDKFGNDIAAFSGAEDVALIEYAKL 129
Db   231 QHVRSTVTELTWGNF---VNSIQTFAGA-----GMLITPCSKGGDCITQSTRURL 279
QY   130 TGRPRFVFLSDTGRNLNPETYQLFDVAVEKHGIRIEYMFPPDAVEVQA-----LVRSKGLF 183
Db   280 SGMFLP-LNSLDA-----KEIYLEVIHNLDPFELLSANTLEDRLAHRWLL 323
QY   184 SFYBDGHECCRVKVRKPLRRALKGLRAWITQKQKDSQPSGTRSRIPVQVQDPA-FEGM-- 240
Db   324 -FFHFGKNVNSNDPELKKLTKLLKNDHIQVGRFDCSSAPDICSNLVYFQPSLAVFKGQGT 382
QY   241 ---DGGIGSLVKKNPVA---NVKGHDIWNPLRTMNPVNSLH---AKGYVSIIGCEPCTRP 291
Db   383 KEYEIIHHGKILYDILAFAKESVNSH--VTTLGPQNFPAKDKEPWLVDFFAPWCPPC-RA 439
QY   292 VLPQHEREGRWWEEDAK--AKECGLHKG-----NVKQKKEEDV-NGNGLSQSHANGDAT 343
Db   440 LLP-ELRRASNLLYGQLKFGTLDTVHEGLCNMNYIOAYPTTVVFNQSNIIHEYEGHSAE 498
QY   344 T----VPDIFNSPNVNLRSRTGIENLAKLEDRKEPWLVLVYAPWCPYCOAMEESYVDLAD 399
Db   499 QILFEFIEDLNN-PSVSLITPTTFNELVTQRKHNEVMVVDYFSPWCHPCQVLPMEKRMAR 557
```







Db 87 IKFYAPRGCHCKTAPTWELSKKEFFGLAGVKIARDVCTAERNICSKYSVR--GYPTLL 144  
Qy 436 FFKPHSSRPITIKYPSKRDVDLSMAFV 462  
Db 145 LF---RGKKVSEHSGRDLDLSLRFV 168

## RESULT 10

US-10-878-556A-1  
; Sequence 1, Application US/10878556A  
; Publication No. US20050266399A1  
; GENERAL INFORMATION:  
; APPLICANT: Hoffmann La-Roche Inc.  
; TITLE OF INVENTION: HCV regulated protein expression  
; FILE REFERENCE: 21762  
; CURRENT APPLICATION NUMBER: US/10/878,556A  
; CURRENT FILING DATE: 2004-06-28  
; NUMBER OF SEQ ID NOS: 199  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 1  
; LENGTH: 414  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (307)..(307)  
; OTHER INFORMATION: Xaa is an undetermined amino acid residue  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: hseup:006101-3-0  
; DATABASE ENTRY DATE: 2003-02-16  
US-10-878-556A-1

Query Match 4.2%; Score 101.5; DB 6; Length 414;  
Best Local Similarity 22.6%; Pred. No. 0.24; Mismatches 34; Indels 47; Gaps 11;  
Matches 50; Conservative 34;  
Qy 250 WNPVANYKGHDIMFLRTMNPVNSLHAKGVSVIGCE-----PCTRPVLPQGHREGR 302  
Db 119 WNDLIG----DKYNSMEDAKVYVAKVDTAHSV--CSAQGVGYPTLKLKFKPQGEA---- 168  
Qy 303 WWEDAKAKECGLHKNVKKQKEDVN--GNLSQSHANGDATVDPDFNSPNVNLRT 360  
Db 169 -----VKYQGRDFQTLNWMMLQTLNEEPVTPPEPV--BPPSAPEL-KQ 209  
Qy 361 GIENLA----KLEDRKEPWLVLVYAPWCPYCOAMEESYVDLADKLACGTGKVKCFRAGD 416  
Db 210 GLYELASNFELHVAQGDHFIFKFFAPWCGHCKALAPTWEQALGLEHSETVKIGK--VDC 267  
Qy 417 EQKEFAKSELQGSFPTLFFPKHSSRPITIKYPSKRDVDLS 457  
Db 268 TQHYELCSGNQVRGYPTLLWF--RDGKKVDQYKG-KRDLES 305

## RESULT 11

US-11-024-959-352  
; Sequence 352, Application US/11024959  
; Publication No. US20060010516A1  
; GENERAL INFORMATION:  
; APPLICANT: FORSTER, RICHARD L.  
; APPLICANT: CONNETT, MARIE B.  
; APPLICANT: EMERSON, SARAH JANE  
; APPLICANT: GRIGOR, MURRAY ROBERT  
; APPLICANT: HIGGINS, COLLEEN M.  
; APPLICANT: LUND, STEVEN TROY  
; APPLICANT: MAGUSIN, ANDREAS  
; APPLICANT: KODRZYCKI, BOB  
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS  
; FILE REFERENCE: 044463-0360  
; CURRENT APPLICATION NUMBER: US/11/024,959  
; CURRENT FILING DATE: 2004-12-30  
; PRIOR APPLICATION NUMBER: 60/533,036  
; PRIOR FILING DATE: 2003-12-30

; NUMBER OF SEQ ID NOS: 782  
; SOFTWARE: Patent in version 3.3  
; SEQ ID NO 352  
; LENGTH: 875  
; TYPE: PRT  
; ORGANISM: Eucalyptus sp.  
US-11-024-959-352

## Query Match 4.0%; Score 99; DB 7; Length 875;

Best Local Similarity 22.0%; Pred. No. 1.2;  
Matches 100; Conservative 52; Mismatches 147; Indels 156; Gaps 27;  
Qy 112 AIAFSGAEDVAL-----IEYAK-----LTC-----RPPRVF-----SLDT 141  
Db 63 SVAFDSAEVLVLGASSGVIKLWDVEAKMVRGLTGHRSNCTAMEFHPFGEFASGSTD 122  
Qy 142 GRLNPETYQLFDAVEK-----HYGRIEYMPD-----AVEVQALVRSKG 181  
Db 123 -----NLKIWDIRKKGCIHTYKGTGTGISTIRFSPDGRVWVSGNDNVVWVWDLTAGKL 176  
Qy 182 L--FSFYEDGHQECRCVRKV--RPLRRAL-----KGLRAW-----ITQRKDQSPGTR 225  
Db 177 LHDQKPHEN-----HIRSIDFHLFELLATGSAADRTVKFWDLETFLIGSSRPEAGVR 230  
Qy 226 SEIPVVOVDP-----AFEGMDGGIGSLVKMNPVANYKHDI--WNLRTMNPVNSLHAKGY 280  
Db 231 A-----IAFHPDGRTLFCGLEDSL--KVYSWEFVICHGVDWGMWSTLADLCIHGDKL----- 280  
Qy 281 VSGCEPCTRPVLPQGHREGRWWEDAKAKECGLHKNVKKQKEDVNGLSQSHANG 340  
Db 281 --LGC-----SYQSGVGVVADASLIE--PYGTNVKPKQ--QKDSGDEIEHQSRP 326  
Qy 341 DA---TTV-----PDIFN-----SPNVNLSRTGIENLAK-----LDRK 372  
Db 327 SAVGTITIRSTIMRCASPDYETKDKNIYVDTASGNPVSSQVRVTFNFAKVTQPLDFND 386  
Qy 373 EPMVLVLYAPWCPYCOAMEESYVDLADKLACGTGK-----VGKFRADGEQKEFAKSE 425  
Db 387 TPNLTL-----RRQGLVTETPDGLSGHVPSKSIQPKVVSRRSDPDGKDSRRRESI 436  
Qy 426 LQGSSEPTILFFPKHSSRPITIKYPSKRDVDLSMA 460  
Db 437 TFSRTKPGMLLRPAHSRRPS-----STKYDVDRLSA 467

## RESULT 12

US-10-467-962B-6  
; Sequence 6, Application US/10467962B  
; Publication No. US20050246784A1  
; GENERAL INFORMATION:  
; APPLICANT: Pleach, Gunnar  
; APPLICANT: Blau, Astrid  
; APPLICANT: Daeschner, Klaus  
; APPLICANT: Klein, Mathieu  
; TITLE OF INVENTION: Identification of Herbicidally Active Substances  
; FILE REFERENCE: 2000 857  
; CURRENT APPLICATION NUMBER: US/10/467,962B  
; CURRENT FILING DATE: 2003-08-14  
; PRIOR APPLICATION NUMBER: PCT/EP02/01466  
; PRIOR FILING DATE: 2002-02-13  
; NUMBER OF SEQ ID NOS: 109  
; SOFTWARE: Patent in Vers. 2.0  
; SEQ ID NO 6  
; LENGTH: 183  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-10-467-962B-6

Query Match 3.7%; Score 91.5; DB 6; Length 183;

Best Local Similarity 31.9%; Pred. No. 0.57; Mismatches 14; Indels 5; Gaps 4;  
Matches 30; Conservative 14;

Qy 354 VVNLSTGIENLAKLEDRKEPWLVLVYAPWCPYCOAMEESYVDLADKLACGTGKVKGR 413

Db 77 VKLSAQELQELVK-GRKVLPLVDYATWCGPCILMAQBLEMLA--VEYESNAIIKVD 133

Qy 414 ADGEQKEFAKSELQSGFPTLFPKHSRPTIK 447

Db 134 TDBEY-EFAR-QMVGRLPTLFFISPDPSKDAIR 165

RESULT 13

US-10-763-712A-48

Sequence 48, Application US/10763712A

Publication No. US20050266541A1

GENERAL INFORMATION:

APPLICANT: Solazyme, Inc.

APPLICANT: Dillon, Harrison F.

TITLE OF INVENTION: Methods and Compositions for Evolving Microbial Hydrogen

TITLE OF INVENTION: Production

FILE REFERENCE: H2042101-CIP

CURRENT APPLICATION NUMBER: US/10/763,712A

CURRENT FILING DATE: 2004-01-21

PRIOR APPLICATION NUMBER: US 10/287,750

PRIOR FILING DATE: 2002-11-04

PRIOR APPLICATION NUMBER: US 10/411,910

PRIOR FILING DATE: 2003-04-12

PRIOR APPLICATION NUMBER: US 60/500,032

PRIOR FILING DATE: 2003-09-03

NUMBER OF SEQ ID NOS: 184

SOFTWARE: Patentin version 3.2

SEQ ID NO 48

LENGTH: 844

TYPE: PRT

ORGANISM: Desulfitobacterium hafniense

US-10-763-712A-48

Query Match 3.7%; Score 91.5; DB 6; Length 844;

Best Local Similarity 19.7%; Pred. No. 5.4;

Matches 84; Conservative 52; Mismatches 142; Indels 149; Gaps 18;

Qy 1 MALAFTSSISAPTSFTFSSSEPKLPQIGSIRISERPIGAVNFNLSQRRSLVKPVNAEPPR 60

Db 240 LSAAYSAIQHVDVTVERHPR-----AGGMRYGIPYRLPKETLDRE--- 283

Qy 61 KDSIVPLAA-----TTIVASASETKEEDPEQIASDLD-----NASPLEIMDRALD 105

Db 284 ----IGLTADLGVKIMTKNALGTHIRLEDLHQ-----DFDANVLAIGSWRATPLQIEGDNLE 336

Qy 106 -----KFGNDIAIAFSG--ABDVALIEYAKLTGRPFVFSLDTGRLN 145

Db 337 GWLGINFLEQVTKGADIKLGEHVVVIGGGNTAIDCARTALKAGSVKLVYRRTREMP 396

Qy 146 PETYQLFADVEKHGIRIEYMPDPDAVEQVALVRSKGLFSFYEDGHQECRCVKVRPLRA 205

Db 397 AESYEVEBEAI--HEGVEMYFL---TAPHKIVABGG-----RKLHLC 432

Qy 206 LKGLRAWITGQRKQDSQCTSEIPVQVDPAPFG--MDGGIG-----SLVK 249

Db 433 IK-----MTLGEPRDS--GRRPPIEGSETAFEDATTIIGAIGQSTNTQFLVHDLPVKLNK 486

Qy 250 WNPVANVKGHDINFLRTMNVNLSHAKGVYSIGCEPCTRPVLPFGQHEREGRWWEDAK 309

Db 487 WGTI-EING-----KTMQTSNNIPAGGDCVTGPATVIOAAGRA-----AE 529

Qy 310 AKGCLLHGNVYKQKEEDVNGLSQSHANGDATVTPDIFNSPVNVLNRTGIENLAKLE 369

Db 530 AMSDFLMKGVYKEQ-----PMDYSCSRGSLDLPQWE 561

Qy 370 DRKEPWL 376

Db 562 FEKIPRL 568

RESULT 14

US-11-156-084-227

Sequence 227, Application US/11156084

Publication No. US20060010515A1

GENERAL INFORMATION:

APPLICANT: Monsanto Technology LLC

TITLE OF INVENTION: Controlled expression of cytokinin biosynthetic genes leads to

TITLE OF INVENTION: agronomically interesting phenotypes

FILE REFERENCE: (38-21)

CURRENT APPLICATION NUMBER: US/11/156,084

CURRENT FILING DATE: 2005-06-17

NUMBER OF SEQ ID NOS: 364

SOFTWARE: Patentin version 3.2

SEQ ID NO 227

LENGTH: 310

TYPE: PRT

ORGANISM: Clostridium perfringens

US-11-156-084-227

Query Match 3.7%; Score 90; DB 7; Length 310;

Best Local Similarity 23.2%; Pred. No. 1.7;

Matches 59; Conservative 46; Mismatches 81; Indels 68; Gaps 16;

Qy 22 KLPQIGSIRISERPIGG-----AVNFNLSQRRSL-----VKPVNA--EPKRDSDI 64

Db 41 KYMDIGSAKISKEEMGGIPHYLIDFVDPSEKFSVAEFKDLTTEKIKDIQSRGKLPILVGG 100

Qy 65 VPLAATTIVAS--ASETKEEDF-----EQIASDLDNASPLEIM-DRALDKFGNDIAIAPS 116

Db 101 TGLYINSIICNMNPAESDKDEEYRELEKIANEHNHLYLHMLXIDLESYN---SIHFN 157

Qy 117 GAEDV-ALIEYAKLTGRPFVFSLDTG-----RLNPETYQLFDAVEKHGIRIE 164

Db 158 NRKRVIRALETYKLTGKPFSSFRKAKNSIYETPNYIVVYVLMMDRAKLYDRINK---RVD 213

Qy 165 YMPFDVAV--EVQALVRSKGLFS-----FYEDGH---QSCCRVKVRPLRA 205

Db 214 IMFEKGLLEEVKNL-KAMGLTDDMQSMKGICYKEVLYLDGKISLEQCIEMIK-QGSRNY 271

Qy 206 LKGLRAWITGQRKD 219

Db 272 AKRQLTWTF---RKD 282

RESULT 15

US-10-858-730-76

Sequence 76, Application US/10858730

Publication No. US2005025568A1

GENERAL INFORMATION:

APPLICANT: Blomquist, Richard B.

APPLICANT: Doten, Reed

APPLICANT: Driggers, Edward M.

APPLICANT: Madden, Kevin T.

APPLICANT: O'Leary, Jessica

APPLICANT: O'Toole, George

APPLICANT: Trueheart, Joshua

APPLICANT: Walbridge, Michael J.

APPLICANT: Yorgey, Peter S.

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID

TITLE OF INVENTION: PRODUCTION

FILE REFERENCE: 14184-030001

CURRENT APPLICATION NUMBER: US/10/858,730

CURRENT FILING DATE: 2004-06-01

PRIOR APPLICATION NUMBER: US 60/475,000

PRIOR FILING DATE: 2003-05-30

PRIOR APPLICATION NUMBER: US 60/551,860

PRIOR FILING DATE: 2004-03-10

NUMBER OF SEQ ID NOS: 364

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 76

LENGTH: 760

TYPE: PRT

ORGANISM: Mycobacterium leprae

US-10-858-730-76

Query Match 3.7%; Score 89.5; DB 6; Length 760;  
Best Local Similarity 19.3%; Pred. No. 7;  
Matches 110; Conservative 76; Mismatches 212; Indels 173; Gaps 25;

Qy	13	TSTPPSEPKLPQIGSIRISERPIGG--AVNFNLQRRSLVKPVNAEPPRKDSIVFLAAT	70
Db	7	TQSFTATVTGSPRIGRRRLKRGATEGYWAKRTSRSELESV-----ASTLRDMWSDIAA-	60
Qy	71	TIVASASETKEEPEQIATDNASPLEIMDRALDKFGNDIAIATFS---GAEDVVALEYA	127
Db	61	---AGLDSVPWTFSYDDQMLDTAFMLGALPARVAQVSDDDLDQYFALARGNNDIKPLEMT	117
Qy	128	KLTC-----RPRVFSLDTGRNLNPETYQLFDAVEKHYGIRIEYMFPPDAVEVQALV	177
Db	118	KWEDTNVHYLVPEIEPATTFSLNPKILGE--LKEALEQRIPSRPVIIGP---VTFL	170
Qy	178	RSKGLFSFYEDGHEQECRVRKVRRLRALGLRA-----WITGQR-----KQSP-----	222
Db	171	LSKGI-----NGGGAPIQRLBELVGIYCTLLSLLAENGARWVQFDEPALVTDLSDPAPALA	226
Qy	223	-----GTRSEIPVVOVDPAFEGMDGGIGSLVKWNEV-----	253
Db	227	EAVYTAGSVSKRPATYVATYFGNPGASLAGLAR-TPIEAIGVDFVCGADTSVAAPPELA	285
Qy	254	-----ANVKGHDIW-----NFLRTMNPVN-----SLH	276
Db	286	GKTLVAGIVDGRNIWRTDLBSALSKLATLIGSAATVAVSTSCSTLHVPSLEPETDLDDN	345
Qy	277	AGYVSIQCEPCTRPVLPQOHEREGRWWDKAKKECG-----LHKGNVKQ	322
Db	346	LRSLAFGAEKVAEVVVLARALRDGR----DAVADEIAASNAAVASRRSDPRLHNGQVRA	401
Qy	323	QKEEDV-----NGNLSQSHANGDA-----TTVPDIFNSPNVNLSTGIENTAKLED	370
Db	402	RIDSIVASGTHRGDAQRRTSQDARLHLPLPTTTIGSYPTQSAIRKARAALQD-AEIDE	460
Qy	371	RKEPWLVLVYAPWCPYQAMEESYVDLADKLAGSTGMKV---GKFRADGQKEFAKSELQ	427
Db	461	AE-----YISRMKKEVAD-AIKLQSLGLDLVLHGEPERNDMVQYFAE---Q	503
Qy	428	LGSF-----PTILFFPKHSSRPTIKYPSEKR	453
Db	504	LCGGFFATQNGWQVSGSRCVRRPPILYGDVSR	534

Search completed: February 18, 2006, 14:52:16  
Job time : 18 secs

**THIS PAGE LEFT BLANK**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n-model

Run on: February 20, 2006, 22:02:50 ; Search time 7245 Seconds  
(without alignments)  
3656.183 Million cell updates/sec

Title: US-10-731-525-8

Perfect score: 2445

Sequence: 1 MALAFTSSISAPTSFTPSSE.....KYPSEKRDVDSLMFVNALR 466

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 2842125653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/abs/ABSSWEB.spool/US10731525/runat.17022006.164857.82/app\_query.fasta.1  
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=bl6sum62 -TRANS=human40.cdi -LIST=45  
-DOCLALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HCP=abes05h  
-USER=US10731525 @CGN 1.1 4939 @runat.17022006.164857.82 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.in.\*

3: gb.env.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pr.\*

9: gb.ro.\*

10: gb.sts.\*

11: gb.sy.\*

12: gb.un.\*

13: gb.vi.\*

14: gb.htg.\*

15: gb.pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description
1	2445	100.0	1629	6 AR533623
2	1951	79.8	1705	15 AF452450
3	1923	78.7	1795	6 AR533622

4	1883.5	77.0	1648	15	AY353089	AY353089 Populus x
5	1832.5	74.9	1386	15	AY568717	AY568717 Lycopersi
6	1832.5	74.9	1835	15	BT013254	BT013254 Lycopersi
7	1829	74.8	1753	15	CRU63784	U63784 Catharanthu
8	1811.5	74.1	1398	6	CQ805882	CQ805882 Sequence
9	1811.5	74.1	1398	6	AX412862	AX412862 Sequence
10	1811.5	74.1	1398	6	AX505803	AX505803 Sequence
11	1811.5	74.1	1398	6	AX651565	AX651565 Sequence
12	1811.5	74.1	1398	15	BT002612	BT002612 Arabidops
13	1811.5	74.1	1710	15	AF424582	AF424582 Arabidops
14	1808	73.9	1608	15	SCJ001208	AJ001208 Brassica
15	1806.5	73.9	1714	15	AY088011	AY088011 Arabidops
16	1801	73.7	1702	15	AY087591	AY087591 Arabidops
17	1799	73.6	103129	15	AP004494	AP004494 Lotus cor
18	1797	73.5	1377	15	AY103313	AY103313 Arabidops
19	1797	73.5	1555	15	AY093319	AY093319 Arabidops
20	1797	73.5	1654	15	AY062665	AY062665 Arabidops
21	1797	73.5	1709	15	AF428445	AF428445 Arabidops
22	1797	73.5	1712	15	AY054175	AY054175 Arabidops
23	1795.5	73.4	1748	15	ATU53864	U53864 Arabidops
24	1784	73.0	1763	15	ATU53865	U53865 Arabidops
25	1758.5	71.9	1687	15	AY088665	AY088665 Arabidops
26	1758.5	71.9	1712	15	ATU96045	U96045 Arabidops
27	1745.5	71.4	1945	15	AF016282	AF016282 Arabidops
28	1745.5	71.4	47319	6	AX059533	AX059533 Sequence
29	1745.5	71.4	103632	15	AF074021	AF074021 Arabidops
30	1745.5	71.4	200001	15	ATCHRIV13	AL161501 Arabidops
31	1738	71.1	1365	6	CQ805458	CQ805458 Sequence
32	1738	71.1	1365	6	AX412892	AX412892 Sequence
33	1738	71.1	1365	6	AX412893	AX412893 Sequence
34	1738	71.1	1365	6	AX507610	AX507610 Sequence
35	1738	71.1	1365	15	AY040005	AY040005 Arabidops
36	1738	71.1	1783	15	AF360192	AF360192 Arabidops
37	1735	71.0	1365	15	AF023167	AF023167 Arabidops
38	1734	70.9	1634	15	ATU43412	U43412 Arabidops
39	1728.5	70.7	1738	15	ATU53866	U53866 Arabidops
40	1717	70.2	1482	15	ATU56922	U56922 Arabidops
41	1715	70.1	1700	15	SCJ001207	AJ001207 Brassica
42	1690.5	69.1	2945	15	AF016284	AF016284 Arabidops
43	1687	69.0	103270	15	ATF1N20	AL022140 Arabidops
44	1687	69.0	196247	15	ATCHRIV56	AL161556 Arabidops
45	1668.5	68.2	1809	15	AY739296	AY739296 Zea mays

ALIGNMENTS

RESULT 1	AR533623	Sequence 7 from patent US 6730827.	1629 bp	DNA	linear	PAT 08-OCT-2004
LOCUS	AR533623	Sequence 7 from patent US 6730827.	1629 bp	DNA	linear	PAT 08-OCT-2004
DEFINITION	AR533623	Sequence 7 from patent US 6730827.	1629 bp	DNA	linear	PAT 08-OCT-2004
ACCESSION	AR533623	Sequence 7 from patent US 6730827.	1629 bp	DNA	linear	PAT 08-OCT-2004
VERSION	AR533623.1	GI:539233592	1629 bp	DNA	linear	PAT 08-OCT-2004
KEYWORDS	AR533623.1	GI:539233592	1629 bp	DNA	linear	PAT 08-OCT-2004
SOURCE	Unknown.	Unknown.	1629 bp	DNA	linear	PAT 08-OCT-2004
ORGANISM	Unknown.	Unknown.	1629 bp	DNA	linear	PAT 08-OCT-2004
REFERENCE	1 (bases 1 to 1629)	Unknown.	1629 bp	DNA	linear	PAT 08-OCT-2004
AUTHORS	Falco, S.C.	Unknown.	1629 bp	DNA	linear	PAT 08-OCT-2004
TITLE	Genes encoding plant adenosine 5'-phosphosulfate reductase	Unknown.	1629 bp	DNA	linear	PAT 08-OCT-2004
JOURNAL	Patent: US 6730827-A 7 04-MAY-2004;	Unknown.	1629 bp	DNA	linear	PAT 08-OCT-2004
FEATURES	E. I. du Pont de Nemours and Company; Wilmington, DE	Unknown.	1629 bp	DNA	linear	PAT 08-OCT-2004
source	Location/Qualifiers	Unknown.	1629 bp	DNA	linear	PAT 08-OCT-2004
ORIGIN	1..1629	Unknown.	1629 bp	DNA	linear	PAT 08-OCT-2004
	/organism="unknown"	Unknown.	1629 bp	DNA	linear	PAT 08-OCT-2004
	/mol_type="genomic DNA"	Unknown.	1629 bp	DNA	linear	PAT 08-OCT-2004
Alignment Scores:	6.49e-175	Length:	1629			
Pred. No.:	2445.00	Matches:	466			
Score:	100.0%	Conservative:	0			
Percent Similarity:	100.0%	Mismatches:	0			
Best Local Similarity:	100.0%	Indels:	0			
Query Match:	100.0%					

DB:	6	Gaps:	0
US-10-731-525-8 (1-466) x AR533623 (1-1629)			
Qy	1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSerGlu	20	
Db	32 ATGGCCCTCGCTTTCACCTTCTCAATTTTCGGACCAACTTCCACTTCCCATCATCGAA	91	
Qy	21 ProLysLeuProGlnIleGlySerIleArgIleSerGluArgProIleGlyGlyAlaVal	40	
Db	92 CCCAAACTTCGGCAAAATGGGTCAATTAGGATTCGGAGAGGCCCATTCGGAGGCGCGTT	151	
Qy	41 AsnPheAsnLeuSerGlnArgArgSerLeuValLysProValAsnAlaGluProProArg	60	
Db	152 AATTTCAATTTATCTCAAGACGGAGCTTGTAAGCCCGTTAAACGCCGAACCTCCACGC	211	
Qy	61 LysAspSerIleValProLeuAlaAlaThrThrIleValalaserAlaSerGluThrLys	80	
Db	212 AAGGATTCCATGTGTTCTTCGGACGCAACCACTGCTTGTCTTCTGAGACGAAA	271	
Qy	81 GluGluAspPheGluGlnIleAlaSerAspLeuAspAsnAlaSerProLeuGluIleMet	100	
Db	272 GAGGAAGATTTGACAGATAGGCAGTGATCTCGACAACTGTTCACTCTTGAAATCATG	331	
Qy	101 AspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSerGlyAlaGluAsp	120	
Db	332 GATAGAGCCCTCGACAAATTCGGCAACGACATAGCTATTGCTTTCAGTCGTCTGAAGAT	391	
Qy	121 ValAlaLeuIleGluThrAlaLysLeuThrGlyArgProPheArgValPheSerLeuAsp	140	
Db	392 GTTGCTTTGATGATGAGTAATTCAGCGGTGACCTTTAGGCTTTTCAGTTTGGAC	451	
Qy	141 ThrGlyArgLeuAsnProGluThrTyrglnLeuPheAspAlaValGluLysHisTyrgly	160	
Db	452 ACTGGAGACTGAACCCAGAAACTTATCACTTTTTCGCGTTGAGAAGCATATTCGA	511	
Qy	161 IleArgIleGluTyrgMetPheProAspAlaValGluValGlnAlaLeuValArgSerLys	180	
Db	512 ATTCGCATTGAGTACATGTTCCCTGATGCTGTTGAGGTTTCAGGCATTTGGTGAGGTAAG	571	
Qy	181 GlyLeuPheSerPheTyrgLysAspGlyHisGlnGluCysAsArgValArgLysValArg	200	
Db	572 GGGTATTCTCTTCTACGAGGATGGGCAACAGAGTGTTCAGGCTGAGAAAGGTGAGG	631	
Qy	201 ProLeuArgArgAlaLeuLysGlyLeuArgAlaTrpIleThrGlyGlnArgLysAspGln	220	
Db	632 CCTTTAAGAGGGGCCCTTTAAGGGTCTCAGAGCATGGATACTGTGTAGAGGAAGACAG	691	
Qy	221 SerProGlyThrArgSerGluIleProValValGlnValAspProAlaPheGluGlyMet	240	
Db	692 TCACCTGGTACTAGTCTGAAATACCGGTGTTTCAGGTTGATCCGGCTTTTGAGGGAATG	751	
Qy	241 AspGlyGlyIleGlySerLeuValLysTrpAsnProValAlaAsnValLysGlyHisAsp	260	
Db	752 GATGGTGAATTTGGAAGCTTGCTGAAGTGGCAACCTGTTGCAAAATGTGAAGGGCCATGC	811	
Qy	261 IleTrpAsnPheLeuArgThrMetAsnValProValAsnSerLeuHisAlaLysGlyTyrg	280	
Db	812 ATATGGAACTTCTTTAGGACCATGAATGTGCTGTGAATTCCTTGTGCATGCAAAAGGATAT	871	
Qy	281 ValSerIleGlyCysGluProCysThrArgProValLeuProGlyGlnHisGluArgGlu	300	
Db	872 GTTTCCATTGGGTGTAGCCCTGCACTAGGCTCTGTTTACCTGGGCAACATGAAAGGAA	931	
Qy	301 GlyArgTrpTrpTrpGluAspAlaLysGluCysGlyLeuHisLysGlyAsnVal	320	
Db	932 GGGAGGTGGTGGTGGAGATGCCAAAGCTAAGGAATGTGCTCTTCAAAAGGAAATGTA	991	
Qy	321 LysGlnGlnLysGluGluAspValAsnGlyAsnGlyLeuSerGlnSerHisAlaAsnGly	340	
Db	992 AAGCAGCAGAAGAGGAGGATGTTAATGGAAATGGGCTATCCCAATCCCATGCAAAATGGT	1051	
Qy	341 AspAlaThrThrValProAspIlePheAsnSerProAsnValValAsnLeuSerArgThr	360	



```
Pred. No.: 1.21e-137 Length: 1705
Score: 1951.00 Matches: 382
Percent Similarity: 88.2% Conservativity: 38
Best Local Similarity: 80.3% Mismatches: 40
Query Match: 79.8% Indels: 16
DB: 15 Gaps: 7

US-10-731-525-8 (1-466) x AF452450 (1-1705)

QY 1 MetAlaLeuAlaPhe-----ThrSerSerIleSerAlaProThrSerThr 15
DB 16 ATGGCTCTCGCGTTTCCACTACTTCTCTCTCAGCAGCTGCTGAGCTTCTAGTCC 75
QY 16 Phe-----ProSerSerGluProLysLeuProGlnIleGlySerIleArg 30
DB 76 TTCTTCTCTGCTCGGATCTTCTCGGACGCAAAAGCTCTCAATCGGTTCGTTCCG 135
QY 31 IleSerGluArgProIleGlyAlaValAsnPheAsnLeuSerGlnArgSerLeu 50
DB 136 TTCCGAGAGGTCGCTTGTTCGTCGTTGTTGTCATGTAACCTCAACGACGCTCTTG 195
QY 51 ValLysProValAsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThr 70
DB 196 GTGAGCCACTCAACGCCGACCGCAACGACGATTCATTTGTTCTCTTTCGACGA 252
QY 71 ThrIleValAlaSerAlaSerGluThrLysGluGluAspPheGluGlnIleAlaSerAsp 90
DB 253 ACTATCGTCTCTCGAGTTGAGAGGAGAGAGAGAGATTTTGGACAAATAGCGAAAGAC 312
QY 91 LeuAspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAsp 110
DB 313 CTTGAAATCATCTCTCTTGAATTTATGATAGGAGGCTCGAGAAATTTGGGAACGAC 372
QY 111 IleAlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThr 130
DB 373 ATCGTATTGCTATTGCTGCTGAGATGTTCTTTGATTGAGTATGCACATTTGACG 432
QY 131 GlyArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGln 150
DB 433 GGTGACCCCTACAGAGTGTATTAGTCTTGACACTGGGAGACTGAATCCAGAAACCTACAA 492
QY 151 LeuPheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAla 170
DB 493 TTTTGTGATCGGTTGAGAGCAATTTGGAATTCACATTTGAGTACATGTTCCCTGATGCG 552
QY 171 ValGluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHis 190
DB 553 GTTGAGGTTGAGGATTTAGTAAAGACTAAAGGCTCTTCTCAATTTACGAGGATGGCAT 612
QY 191 GlnGluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArg 210
DB 613 CAAGAGTGTCTCGAGTAAGAGAGTGGAGCCCTTGAGGAGAGCCCTTAAGGCTCTCAA 672
QY 211 AlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProVal 230
DB 673 GCATGGATCACTGGACAGAGAAAGATCAGTCTCTGCTAGTACTAGGCTCGAAATCCCTGTT 732
QY 231 ValGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrp 250
DB 733 GTCAGGTTGATCCTGTTTGGAGGACTGGATGGTGGAAATGGCAGCCTGGTGAAGTGG 792
QY 251 AsnProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnVal 270
DB 793 AACCCGGTGGCAATGTAATGCTCTAGACATATGAGCTTCTTAGGACCATGGATGTT 852
QY 271 ProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArg 290
DB 853 CCTGTAATTCATTGCAATCCCAAGGATATGTTTCGATTGGCTGTGAGCGCTGCACAAGG 912
QY 291 ProValLeuProGlyGlnHisGluArgGluGlyArgTyrTrpGluAspAlaLysAla 310
DB 913 CCAGTTTTACCGGCTCAACATGAAGAGAGGAAGGAGTGGTGGGAGGATGCCAAGCC 972
```

---

```
QY 311 LysGluCysGlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluAspValAsnGly 330
DB 973 AAGGAGTGTGCTCTTCAAAAGGTAATATAAATCAAGATGCTGCCAGCTTAATGGA 1032
QY 331 AsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsn 350
DB 1033 AATGGGCGCTCCCAA-----GCAAAATGGATCTGCC---ACTGTCTGCTGACATTTTCAAC 1083
QY 351 SerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAsp 370
DB 1084 TCCAGGATGTGTGAGCTGAGCAGGTCTGGGATTCGAAATTTGGCAAAATTTGGAGAAC 1143
QY 371 ArgLysGluProTrpLeuValLeuValLeuTyrAlaProTrpCysProTyrCysGlnAlaMet 390
DB 1144 CGAAAGAACCCCTGGCTTGTGTCTCTATGCACCATGGTGGCGCTCTGTCTGAGGCTATG 1203
QY 391 GluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGly 410
DB 1204 GAGGAATCATATGTTGATTGGCAGAAAGTTAGCTGGGTCA---GGAGTGAAGGTTGCA 1260
QY 411 LysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLysSer 430
DB 1261 AATTTCAGAGCGATGGAGACAGAGAAATATGCAAAAGTGAACCTGAGCTGGAGAGC 1320
QY 431 PheProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSer 450
DB 1321 TTCCCCACAATACTACTCTTCTCCCAACACTCTTCTCAACCA---ATTAAATACCTTCG 1377
QY 451 GluLysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
DB 1378 GAAAGAGAGATGTTGATTCATTGACGGCATTCGTGAATGCTTTACGG 1425

RESULT 3
LOCUS AR533622 1795 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 5 from patent US 6730827.
ACCESSION AR533622
VERSION AR533622.1 GI:53923591
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS Falco,S.C.
TITLE Genes encoding plant adenosine 5'-phosphosulfate reductase
JOURNAL Patent: US 6730827-A 5 04-MAY-2004;
E. I. du Pont de Nemours and Company; Wilmington, DE
FEATURES
source 1..1795
Location/Qualifiers
/mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.: 1.67e-135 Length: 1795
Score: 1923.00 Matches: 378
Percent Similarity: 87.2% Conservativity: 37
Best Local Similarity: 79.4% Mismatches: 45
Query Match: 78.7% Indels: 16
DB: 6 Gaps: 7

US-10-731-525-8 (1-466) x AR533622 (1-1795)
QY 1 MetAlaLeuAlaPheThrSerSerIleSer-----AlaProThrSerThr 15
DB 56 CTTCGGTTCCTCACTACTTCTTCTCTTTCAGCTGACAGCAGCAGCGCTCGAGCTCT 115
QY 16 Phe-----ProSerSerGluProLysLeuProGlnIleGlySerIleArg 30
DB 116 TTCTTCTCGCGCTTGGATCTTTCATCGGAGCTTAAAGCTCCGCAATTTGGTTCCTTCGG 175
QY 31 IleSerGluArgProIleGlyAlaValAsnPheAsnLeuSerGlnArgSerLeu 50
DB 913 CCAGTTTTACCGGCTCAACATGAAGAGAGGAAGTGGTGGGAGGATGCCAAGCC 972
```

```
Db 176 TTTCCGAGAGGCTCAAGTTTCGTCTGCTGTTGTTTAATTAACCAAGACGCTCCTCG 235
Qy 51 ValLysProValAlaAlaGluProProArgLysAspSerIleValProLeuAlaAlaThr 70
Db 236 GTGAGGCACTCAATGCGCAACCGCAACGGAATGATTCTGTGTCTCTTGCAGCA--- 292
Qy 71 ThrIleValAlaSerAlaSerGluThrLysGluGluAspPheGluGluInIleAlaSerAsp 90
Db 293 ACTATGCTTCTCTGAGGTGAGAGGAGAGGAAGATTTTGAGCAATTAGCGAAAGAC 352
Qy 91 LeuAspAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAsp 110
Db 353 CTTGAAATTCATCTCTCTTGAGATTATGATAAGSCCTCGAGAAATTTGGCAACGAC 412
Qy 111 IleAlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluThrAlaLysLeuThr 130
Db 413 ATCGCTATTGCTTGTAGTGTCTGAAGATGTTGCTTTGATTGAGTATGCACATTTGCAGC 472
Qy 131 GlyArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGln 150
Db 473 GGTGACCCCTACAGAGGTGTTAGTCTTGACACTGGGAGACTGAACCCAGAAACCTACAA 532
Qy 151 LeuPheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAla 170
Db 533 TTTTTCAGCGCTGTTGAGAGCATTATGGAATTCATATTGATACATGTTCCCTGATGCG 592
Qy 171 ValGluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHis 190
Db 593 GTTGAGGTTGAGGCATTAGTAAGAACTAAGGGGCTCTCTCATTTTACGAGGATGGGCAT 652
Qy 191 GlnGluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArg 210
Db 653 CAAGAGTGCTGTAGATAAGAAAGGTGAGGCCCTTGAGGAGAGCCCTTAAGGGTCTCAAA 712
Qy 211 AlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProVal 230
Db 713 GCATGGATTACTGCACAGAAAGACACAGCTCTCTGTTACTAGTCTGAAATCCCTATT 772
Qy 231 ValGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrp 250
Db 773 GTCCAGGTTGATCCTGTTTGTAGGGACTGGATGGTGGAAATTTGCGAGCCTGGTGAAGTG 832
Qy 251 AsnProValAlaAlaValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnVal 270
Db 833 AACCCGGTTGCAATGTTAATGGTCTAGACATATGGAACTTCTCTTAGACCATGAATGTT 892
Qy 271 ProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArg 290
Db 893 CTTGTAATTCATTGCTATCCCAAGGATATGTTTCGATTGGCTGTGAGCCATGCACAAGG 952
Qy 291 ProValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpGluAspAlaLysAla 310
Db 953 CCGGTTTTTACCGGACCAACATGAAGAAGAAAGGTGGTGGTGGGAGGATGCCAAAGCC 1012
Qy 311 LysGluCysGlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluAspValAsnGly 330
Db 1013 AAGGAGTGTGCTTTCACAAAGGTAATTTGAAACAGGAAGATGCTGCCACCTAATAGGA 1072
Qy 331 AsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsn 350
Db 1073 AATGGGACCTCCCAA-----GGAATGGCTCTGCC---ACTGTTGTGTGACATTTTCATC 1123
Qy 351 SerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAsp 370
Db 1124 TCCAGAAATGGTGCAGCTTGCAGCAGGTCCGGGATTTGAGAAATTTGGCAAAATTAGAGAAC 1183
Qy 371 ArgLysGluProTrpLeuValIleValThrAlaProTrpCysProTyrCysGlnAlaMet 390
Db 1184 GGAAGAAGAACACTGGCTTGTGTGCTCTATGCACCATGTCGCCGCTCTGTGCAGCCTATG 1243
Qy 391 GluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGly 410
Db 1244 GAGGAGTGTATGTTGATCTCGCAGAGAAGTTAGCAAGGTCA---GGAGTGAAGGTTGCA 1300
```

```
Qy 411 LysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySer 430
Db 1301 AAATTCAGACCCGATGAGAGCAGAGGAATATGCAAAAGAGTGAACCTGCGATTGGGAGC 1360
Qy 431 PheProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSer 450
Db 1361 TTCCCCACAATACTTCTCTTCCCAAGCACTCTTCTCAACCA---ATTAAGTACCCCTCA 1417
Qy 451 GluLysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
Db 1418 GAAAGAGAGAGATGTTGATTCATTGACGGCATTCGTGAATGCCCTTACGG 1465

RESULT 4
LOCUS AY353089 1648 bp mRNA linear PLN 12-JAN-2005
DEFINITION Populus x canescens adenosine 5' phosphosulfate reductase mRNA,
complete cds.
ACCESSION AY353089 GI:34099826
KEYWORDS Populus alba x Populus tremula
SOURCE Populus alba x Populus tremula
ORGANISM Populus alba x Populus tremula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE 1 (bases 1 to 1648)
Kopriva,S., Hartmann,T., Massaro,G., Hoenicke,P. and Rennenberg,H.
Regulation of sulfate assimilation by nitrogen and sulfur nutrition
in poplar trees
TREES (Berl. West) 18, 320-326 (2004)
REFERENCE 2 (bases 1 to 1648)
Hartmann,T., Hoenicke,P., Wirtz,M., Hell,R., Rennenberg,H. and
Kopriva,S.
Regulation of sulphate assimilation by glutathione in poplars
(Populus tremula x P. alba) of wild type and overexpressing
gamma-glutamylcysteine synthetase in the cytosol
J. Exp. Bot. 55 (398), 837-845 (2004)
REFERENCE 3 (bases 1 to 1648)
Kopriva,S.
Direct Submission
Submitted (28-JUL-2003) Institute of Tree Physiology,
Georges-Koehler-Allee 053, Freiburg 79085, Germany
JOURNAL Location/Qualifiers
FEATURES
source
1..1648
/organism="Populus alba x Populus tremula"
/mol_type="mRNA"
/db_xref="taxon:80863"
/clone="INRA 717 1B4"
74..1471
/codon_start=1
/product="adenosine 5' phosphosulfate reductase"
/protein_id="AA057202.1"
/db_xref="GI:34099827"
/taxelation="MALVS5555SISASGFSRSQELKAPQFGSKLLDRPTTFVNVS
SQRCAVKPLNAEPKNGSVPLAATIAPEIAEKVVEDYDEYKALVELVNGSPLEIM
DKALEKGGDDIAIFSGAEDVALIEYAKLTGRPRVFSLDTGRNLNPTIYHFDQVEKH
YGRITYEPDFAVEQALNRKGLFSYEDGHOCCRRKVRKVRLLRALKGLRAWITGG
LHQSGYTSICEPTRPVLPGQHERGRWWEDATACEGLHKGNLKQGDAAQLNGNG
NGRAHGAATVADIENSENLNLSPGIENLLKLENKBPWLVLVAPWCQFCQME
ASVLEADKLKAGSVKVKFRADGQKGFQKQELQSGFPTILFFPKHSSRPKYPSE
KRDVDSLMTFVNALR"

Alignment Scores:
Pred. No.: 1,43e-132 Length: 1648
Score: 1893.50 Matches: 366
Percent Similarity: 85.3% Conservative: 35
Best Local Similarity: 77.9% Mismatches: 60
Query Match: 77.0% Indels: 9
DB: Gaps: 7
```

```
US-10-731-525-8 (1-466) x AY353089 (1-1648)
Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPhe-----ProSer 18
Db 74 ATGCATTTGAGTGTAGTCTTCTCTCGATTTCTGCTTCTGGTTTTCACGTTCTAGT 133
Qy 19 SerGluProLysLeuProGlnIleGlySerIleArgIleSerGluArgProIleGlyGly 38
Db 134 CAAGAACTTAAAGCACCCAGTTTGGTTCTTTTAAGCTGTAGATCGGCGGACACA--ACA 190
Qy 39 AlaValAsnPheAsnLeuSerGlnArgArgSerLeuValLysProValAsnAlaGluPro 58
Db 191 TTTGTTAATGTTCT--TCACAAAGACGCTGCTGCTGTAAGCCCTTTAAATGCTGAGCCT 247
Qy 59 ProArgLysAspSerIleValProLeuAlaAlaThrIleValAlaSerAlaSerGlu 78
Db 248 AAAAGGAATGGTCTGTGTTCTCTCGTGCACATATCGTGTCCAGAGATGCGAGAG 307
Qy 79 -----ThrLysGluGluAspPheGluGlnIleAlaSerAspLeuAspAlaSerPro 96
Db 308 AAAGTAGAGGTGGAGGAAGATTATGAGAAGTTAGCTAAGGAGCTTGTAAATGGTTCTCCT 367
Qy 97 LeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSer 116
Db 368 TTGGAATATTATGGATAAGGCATTGAGAAATTTGGGATGACATTTGCTATTGCTTTTCAGT 427
Qy 117 GlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArgProPheArgVal 136
Db 428 GTGCGCCAGAGATGTGTTGATTGATGATGCAAGTTGACTGGCGCACCGTTCAGGGTG 487
Qy 137 PheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAspAlaValGlu 156
Db 488 TTCTCACTGGATACGGGGCGGTGGAACCCGAGACATACCATTTCTTTTGACCAAGTTGAG 547
Qy 157 LysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaValGluValGlnAlaLeu 176
Db 548 AAGCACTATGCGCATTCGAATTTGATGATGATGCAAGTTGACTGGCGCACCGTTCAGGGTG 607
Qy 177 ValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGluCysAspGVal 196
Db 608 GTAAGCAACAGGGTTTATTTCTCTTTTACAGATGGGATCAGAGTGTGCGCGGTGTC 667
Qy 197 ArgLysValArgProLeuArgAlaLeuLysGlyLeuArgAlaTrpIleThrGlyGln 216
Db 668 AGGAAGGTGAGACCTTTGAGCGGGCTCTGAAGGGCTTCGGGCTCGATCACTGGCCAA 727
Qy 217 ArgLysAspGlnSerProGlyThrArgSerGluIleProValValGlnValAspProAla 236
Db 728 AGGAAGGATCAATCTCCAGGAGCAGAGTCTGAAATTCAGTTGTTCAGGTGGACCCCGTT 787
Qy 237 PheGluGluMetAspGlyIleGlySerLeuValLysTrpAsnProValAlaAsnVal 256
Db 788 TTTGAGGATTTGATGTGGGGCTGGAAGCCTTATCAAGTGGATCAATGGCAATGTT 847
Qy 257 LysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAsnSerLeuHis 276
Db 848 GAGGACCAAGATGATGGAAGTTCTTTCGAACATGATGATGTGCTGTGAACTCATTCGAT 907
Qy 277 AlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeuProGlyGln 296
Db 908 TCAAAGGATACATCTCTATTGTTGCGAGCCTTTCGACCAAGGCCAGTCTTCTTCCTGGCAA 967
Qy 297 HisGluArgGluGlyArgTyrTrpTrpGluAspAlaLysAlaLysGluCysGlyLeuHis 316
Db 968 CATGAGAGAAGACCGTGTGTGGGAGATGCCACAGCTAAGGAATGTGTCCTTCAT 1027
Qy 317 LysGlyAsnValLysGlnGlnLysGluGluAspValAsnGlyAsnGlyLeuSerGlnSer 336
Db 1028 AAAGGAATCTGAACAGGGTGATCGACCCCACTTATGCAATGGAACGGGGCTGCC 1087
Qy 337 HisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSerProAsnValValAsn 356
|||||
```

```
Db 1088 CATGCCAAATGGG--GCTGCCACTGTTGCTGATATTATTTCAACTCCGAGAAATTTGGTCAAT 1144
Qy 357 LeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArgLysGluProTrpLeu 376
Db 1145 TTGAGCAGGCCCGAATCGAGAACTTGTAAATATTAGAGAACCGAAGAACCATGGCTT 1204
Qy 377 ValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMetGluGluSerTyrValAsp 396
Db 1205 GTGGTGTCTACGCCCATGGTGCCTAATTTTGCAGAGTATGGAAGCATCTTATGTTGAA 1264
Qy 397 LeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLysPheArgAlaAspGly 416
Db 1265 TTGGCCGATAAGTTAGCAGGAGT---GGAGTAAAGGTGGGAAATTTTCAGGCGAGATGGC 1321
Qy 417 GluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPheProThrIleLeuPhe 436
Db 1322 GATCAGAGGGGATCTCAAAACCAAGAAATTACAGCTAGGAAGTTTTCACACATCTTTC 1381
Qy 437 PheProLysHisSerSerArgProThrIleLysTyrProSerGluLysArgAspValAsp 456
Db 1382 TTCCTTAAACACTCATCTGCCCA---ATCAAGTACCCCTCTGAGAAAGGAGCGTGGAT 1438
Qy 457 SerLeuMetAlaPheValAsnAlaLeuArg 466
Db 1439 TCATTGATGACTTTTGTGAATGCTCTCCGG 1468

RESULT 5
LOCUS AY568717 1386 bp mRNA linear PLN 24-JUN-2005
DEFINITION Lycopersicon esculentum adenyl-1-sulfate reductase mRNA, complete cds.
ACCESSION AY568717
VERSION AY568717.1 GI:51457939
KEYWORDS
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; Jamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 1386)
AUTHORS Coker, J. S., Vian, A. and Davies, E.
TITLE Identification, accumulation, and functional prediction of novel tomato transcripts systemically upregulated after fire damage
JOURNAL Physiol. Plantarum 124 (3), 311-322 (2005)
REFERENCE 2 (bases 1 to 1386)
AUTHORS Coker, J. S., Vian, A. and Davies, E.
TITLE Direct Submission
JOURNAL Submitted (08-MAR-2004) Botany, North Carolina State University, Gardner Hall, Campus Box 7612, Raleigh, NC 27695, USA
FEATURES
source 1..1386
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
CDS 1..1386
/codon_start=1
/product="adenyl-1-sulfate reductase"
/protein_id="AAU03359.1"
/db_xref="GI:51457940"
/translation="MALTTFTSSAIHGSLSSSSSSYEQPKVSQLGTQFPLDRQLSS
TVNLSRRSAVKPLVYAEPRKNDISVPSAATVVAPEGVESVEAEKELAKLQNASPL
EYHDKALEKFGDDIALAFSGAEDVALIEVAHLTGRVRFVSLDTGRINPETYQLFDTV
EKMGRLRIEYMPDPSVEQALVTKGLFDFEDHQBCCRVKRVPLRRALGLRAWI
TGORKDQSPCTRSEIPIVOVDPSEGLDGGAGSLVKWNPVANDGKDIWNFLAMNVP
VNSLHQGVYSIGCEPCTRPVLPQHERERHWEDEKAKACECLHKGNIKDEVTNGAA
QNGTATVADIPDKTIDIVLSKPVENVKLEDRPMLVLYAPWCQPCQAMEGSTV
ELAEKLAGSGVKFRADGDKQAFQEELQLGSPFTILFFPKHSSKAIKYPSEKRDV
DSLAFVFNALR"
transit_peptide 1..213
/notes="chloroplast transit peptide"
evidence=not_experimental
misc_feature 346..963
/notes="Region: phosphoadenosine phosphosulfate reductase"
```

```
misc_feature      domain"
1048..1377
/notes="Region: thioredoxin domain 2"

ORIGIN
Alignment Scores:
Pred. No.:      8 19e-129      Length:      1386
Score:          1832.50      Matches:      355
Percent Similarity: 85.4%      Conservative: 49
Best Local Similarity: 75.1%      Mismatches: 50
Query Match:      74.9%      Indels:      19
DB:              15          Gaps:         8

US-10-731-525-8 (1-466) x AY568717 (1-1386)
Qy 1 MetAlaLeuAlaPheThrSer-----SerIleSerAlaProThrSer 14
Db 1 ATGGCTTTTGACTTTTCACTTCTTCATCTGCAATTCATGGCTCTTTGTCTTCTTCATCTTCT 60
Qy 15 ThrPheProSerSerGluProLysLeuProGlnIleGlySerIleArgIleSerGluArg 34
Db 61 TCTTAT-----GAACAACCCAAAGTATCCCAATTGGGTACCTTTGAGCCCAATTGGTAGG 114
Qy 35 ProIleGlyGlyAlaValAsnPheAsnLeuSerGlnArgArgSerLeuValLysProVal 54
Db 115 CCTCAACTATTGTCGTCACACTGTTTGNATTTCTCGGAGCGGTCGGCAGTGAAGCCATTG 174
Qy 55 AsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThrIleValAla 74
Db 175 TATGCTGAACCTAAGAGGAATGATCAATAGTTCCGTCAGCAGCTACCATGCTGGCTCCT 234
Qy 75 SerAlaSerGluThrLysGlu--GluAspPheGluGlnIleAlaSerAspLeuAspAsn 93
Db 235 GAGGTAGGAGAGAGTGTGAGCGCAGAGGACTTTGAGAAATTTGGCTAAGGAGCTTCAAAAT 294
Qy 94 AlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAlaIle 113
Db 295 GCTTCCCTCTTTGAGGTATGGAACAAGACACTTGAGAAATTTGGAGATGACATTTGCTATT 354
Qy 114 AlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArgPro 133
Db 355 GCTTTTCAGTGTGCTGAAGATGTTGCTTTGATAGATAGCAGACATTTAACTGGAGACCA 414
Qy 134 PheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAsp 153
Db 415 TACAGAGTATTTCAGCCCTTGATCTACTGGAGGTGTAACCCGGAGACATACCAATTATTGAC 474
Qy 154 AlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaValGluVal 173
Db 475 ACAGTGGAGAACCACTATGGCATTCGCATTAATACATGTTCCCTGATTCACTTGAAGTT 534
Qy 174 GlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGluCys 193
Db 535 CAGCGTTGGTTAGGACCAAGGGCTTTCTTCTTCTATGAGTGGCCACCAAGAGTGT 594
Qy 194 CysArgValArgLysValArgProLeuArgAlaLeuLysGlyLeuArgAlaTyrIle 213
Db 595 TGCCGTGTAAAGAAAGGTGTAGCCCTTTGAGGAGAGCTCTAAAGGGGCTTACGGCGCTGGATC 654
Qy 214 ThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValValGlnVal 233
Db 655 ACAGGCCAGCGTAAGATCAGTCCCTCGGAACCTCGATCAGAAATCCCATTTGTTCAAGTG 714
Qy 234 AspProAlaPheGluGlyMetAspGlyIleGlySerLeuValLysTyrAsnProVal 253
Db 715 GACCCCTTTTGGGGGTGGATGGCGGTGCTGGTAGCTTGGTGAAGTGAACCCCTGTG 774
Qy 254 AlaAsnValLysGlyHisAspIleTyrAsnPheLeuArgThrMetAsnValProValAsn 273
Db 775 GCTAATGTGGACGGAAAGATATTGGAACTTCCTGCGTGCATGAATGTGCTGTTGAAC 834
Qy 274 SerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeu 293
```

```
Db 835 TCATTGCTTTCACAAGGATATGATATCCATTGGATGCGAACCTTGCACAAGGCCAGTTCTTA 894
Qy 294 ProGlyGlnHisGluArgGluGlyArgTyrTrpGluAspAlaLysAlaLysGluCys 313
Db 895 CCAGGGCAACACGAGAGAGAGGGAAGATGTGTGGGAAGATGCCAAGGCCCAAGAGTGT 954
Qy 314 GlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluAspValAsnGlyAsnGlyLeu 333
Db 955 GGCTTGCACAAGGCCAACATC-----AAGGATGAACCTGTAAATGGC----- 996
Qy 334 SerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSerProAsn 353
Db 997 ---GCTGCCCAAAACAAATGGTACTGCT---ACCGTTGCTGATATTTTGTATCCCAAGGAC 1050
Qy 354 ValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArgLysGlu 373
Db 1051 ATTGTTTACCTTGGATGAAGCCTGGAGTGTGAGAACCTAGTAAATTTGGAAGACCCGAAGAG 1110
Qy 374 ProTyrLeuValValLeuTyrAlaProTyrCysGlnAlaMetGluGluSer 393
Db 1111 CCTGGCTCGTGTCTTTATGACCTTGGTGCATTTTGGCCAGGCAATGGAAGGATCC 1170
Qy 394 TyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLysPheArg 413
Db 1171 TATGTTCAATTTGGCTGAGAAAGTTGGCTGGTTCT---GGTGTGAAAGTAGGAAAATTCAAG 1227
Qy 414 AlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPheProThr 433
Db 1228 GCAGATGGTACCAAGAAGCATTTGCACAGAAGATTTGCAGCTTGGCAGCTTCCCTTACA 1287
Qy 434 IleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSerGluLysArg 453
Db 1288 ATACTCTTCTCCCAAGACACTCTTCAAG---GCCATTAAAGTACCTTCCAGAGAAGG 1344
Qy 454 AspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
Db 1345 GACGTAGACTCTTGTGCTGGCTTTTGTGAATGCTCTCAGA 1383

RESULT 6
BT013254      1835 bp      mRNA      linear      PLN 11-MAY-2004
LOCUS      Lycopersicon esculentum clone 134763R, mRNA sequence.
DEFINITION
ACCESSION      BT013254
VERSION      BT013254.1 GI:47104669
KEYWORDS      FLI_CDNA.
SOURCE      Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM      Lycopersicon esculentum
REFERENCE      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
TITLE      asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
JOURNAL      1 (bases 1 to 1835)
Kirkness,E.F., Wang,W. and Vazeille,A.
Direct Submission
Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
Medical Center Drive, Rockville, MD 20850, USA
FEATURES
source
1..1835
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
/clone="134763R"
/tissue_type="callus tissue"
/note="TCADM70"

ORIGIN
Alignment Scores:
Pred. No.:      1.14e-128      Length:      1835
Score:          1832.50      Matches:      355
Percent Similarity: 85.4%      Conservative: 49
Best Local Similarity: 75.1%      Mismatches: 50
Query Match:      74.9%      Indels:      19
DB:              15          Gaps:         8
```

US-10-731-525-8 (1-466) x BT013254 (1-1835)

QY 1 MetAlaLeuAlaPheThrSer-----SerIleSerAlaProThrSer 14  
 Db 93 ATGGCTTTGACTTTTCACTTCTTCATCTGCAATTCATGGCTTTGTCTTCTTCATCTCT 152  
 QY 15 ThrPheProSerSerGluProLysLeuProGlnIleGlySerIleArgIleSerGluArg 34  
 Db 153 TCTTAT-----GAACAACCAAGATATCCCAATTTGGGTACCTTTACGCCATTGGATAGG 206  
 QY 35 ProIleGlyAlaValAlaValAsnLeuSerGlnArgSerLeuValIleProVal 54  
 Db 207 CCTCACTATTGCTGCTCACTGTTTGAATTTCTCGAGGCGTTCGGCAGTGAAGCCATTG 266  
 QY 55 AsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThrThrIleValAla 74  
 Db 267 TATGCTGAACCTAAGAGGAATGATTCAATAGTTCCGTACGACGCTACCATCGTGGCTCCT 326  
 QY 75 SerAlaSerGluThrLysGlu---GluAspPheGluGlnIleAlaSerAspLeuAspAsn 93  
 Db 327 GAGGTAGGAGAGAGTGTGTAGGCACAGAGACTTTTGAGAAATTTGGCTAAGGAGCTTCAAAAT 386  
 QY 94 AlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAlaIle 113  
 Db 387 GCTTCCCTCTTGAAGTTATGGAACAAGCACTTGAGAAATTTGGAGATGACATTGCTATT 446  
 QY 114 AlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArgPro 133  
 Db 447 GCTTTCAGTGGTGTGAAGATGTGCTTTGATAGATGACGACATTTAACTTGACGACCA 506  
 QY 134 PheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAsp 153  
 Db 507 TACAGACTATTACGCTTGATCTAGCGAGGTGTAACCCGAGACATACCAATTATTGAC 566  
 QY 154 AlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaValGluVal 173  
 Db 567 ACAGTGGAGAGCACTATGCATTCGCAATTCGAATACATGTTCCTCGATGATGATGAAGTT 626  
 QY 174 GlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGluCys 193  
 Db 627 CAGCGTGTGTAGGACCAAGGCTTTCTTCTTCATGAGATGGCCACCAAGAGTGT 686  
 QY 194 CysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaTrpIle 213  
 Db 687 TGCGGTGTAAGGAAGTGTAGCCCTTTCAGGAGAGCTCTAAAGGCTTACCGCGCTGGATC 746  
 QY 214 ThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValValGlnVal 233  
 Db 747 ACAGGCCAGCGTAAGATCAGTCCCTCGAACTCGATCAGAAATCCCATTTGTTCAAGTG 806  
 QY 234 AspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsnProVal 253  
 Db 807 GACCTTCTTTGAGGGGTTGGATGGCGGTGCTGTGAGCTTGTGAGTGGAAACCCCTGTG 866  
 QY 254 AlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAsn 273  
 Db 867 GCTAATGTGACGCGAAAGATATTTGGAATCTCTCGGTGCTCCATGATGTCCTGTGAAC 926  
 QY 274 SerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeu 293  
 Db 927 TCATTGCAATTCACAAGGATATGATCCATTGGATGCGAACTTGCACAAAGCCAGTTCTA 986  
 QY 294 ProGlyGlnHisGluArgGluGlyArgTyrTrpTrpGluAspAlaLysAlaLysGluCys 313  
 Db 987 CCAGGGCACCAGAGAGAGGGAAGATGTGTGGGAAGATGCCAAGGCCAAGGAGTGT 1046  
 QY 314 GlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluAspValAsnGlyAsnGlyLeu 333  
 Db 1047 GGCTTGCAAGGCGCAATC-----AAGGATGAACTGTAAATGCG----- 1088  
 QY 334 SerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSerProAsn 353  
 Db 1089 ---GCTGCCCAAAACAAATGGTACTGCT---ACCGTTGCTGATATTTTGTATCAAGGAC 1142

QY 354 ValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArgLysGlu 373  
 Db 1143 ATTGTTACCTTGAAGTAAAGCTGGAGTTGAGAACTAGTAATAATTGGAAAGCCGAGAGAG 1202  
 QY 374 ProTrpLeuValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMetGluGluSer 393  
 Db 1203 CCTGGCTCGTCTTCTTTATGCACCTTGTGCAATTTTGGCCAGGCAATGGAAGGATCC 1262  
 QY 394 TyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLysPheArg 413  
 Db 1263 TATGTTGAATTTGGCTGAGAAAGTTGGCTGGTCT---GGTGTGAAAGTAGGGAATTCAGG 1319  
 QY 414 AlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPheProThr 433  
 Db 1320 GCAGATGGTGACCAAGAACATTTGCACAGAGAAATTCAGCTTGGCAGCTTCCCTTACA 1379  
 QY 434 IleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSerGluLysArg 453  
 Db 1380 ATACTCTTCTTCCCAAAGCACTCTTCAAAG---GCCATTAAAGTACCTTCAGAGAGAGG 1436  
 QY 454 AspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466  
 Db 1437 GACGTAGACTCTTGTGCTGGCTTTTGTGAATGCTCTCAGA 1475

RESULT 7  
 CRU63784  
 LOCUS  
 DEFINITION  
 CRU63784  
 DEFINITION  
 complete cds; nuclear gene for chloroplast product.  
 U63784  
 VERSION  
 U63784.2 GI:12831473  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Catharanthus roseus (Madagascar periwinkle)  
 Catharanthus roseus  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Gentianales; Apocynaceae; Rauvolfioideae; Vinaceae; Catharanthus.  
 REFERENCE  
 1 (bases 1 to 1753)  
 Prior,A., Uhrig,J.F., Heins,L., Wiesmann,A., Lillig,C.H., Stoltze,C., Solli,J. and Schwenn,J.D.  
 Structural and kinetic properties of adenylyl sulfate reductase from Catharanthus roseus cell cultures  
 Biochim. Biophys. Acta 1430 (1), 25-38 (1999)  
 1002930  
 REFERENCE  
 2 (bases 1 to 1753)  
 Uhrig,J.F., Prior,A. and Schwenn,J.D.  
 Direct Submission  
 Submitted (15-JUL-1996) Biochemistry of Plants,  
 Ruhr-Universitaet-Bochum, Universitaetsstr. 150, Bochum 44780,  
 Germany  
 3 (bases 1 to 1753)  
 Prior,A. and Schwenn,J.D.  
 Direct Submission  
 Submitted (03-FEB-2000) Biochemistry of Plants,  
 Ruhr-Universitaet-Bochum, Bochum 44780, Germany  
 REMARK  
 Sequence update by submitter  
 On Feb 15, 2001 this sequence version replaced gi:1488042.  
 FEATURES  
 Location/Qualifiers  
 1..1753  
 /organism="Catharanthus roseus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4058"  
 1..1753  
 /gene="par2neu"  
 /evidence=experimental  
 106..1500  
 /gene="par2neu"  
 /function="reduction of adenylyl sulfate (APS)"  
 /note="C-terminus similar to pdi and beta-subunit of 4-hydroxylase, PDB Accession Number: 1MEK"  
 /codon\_start=1

```
/evidence=experimental
/product="pAPS-reductase-like protein"
/protein_id="AAB05871.2"
/db_xref="GI:12831474"
/translation="MALAFTSSTAIHGSLSSFEQTKAAQAQFSGFQPLDRPHITPS
VNYRRRLAVKPNIAEPKRNESIVPSAATTVAPEBEKVDEYDEKLDELONASPLE
IMDKSLAKGNDIAIAPSGAEDVALIEYAHLTGRPFVSLDTGRINLPETYKFFDVTVE
KOYGIHTEYMFDPDAVEQALVRSKGLFSFYEDGHOECCKVRKVPRLRALKLRAMIT
QQRKDQSPGRSEIPVQVDPVPEGMGGVSLVKNPVPANVSGKDIWNFLRAMDVPV
NTHLSQGYVSIQEPCTRPVLPQGHREGRWMDAKAKCEGLHGKNIKKEETLNNGN
GAVNGSDTIADIFDNNVTSLSRPGIENLLKEERREAWLVLYAPWCRFOAMEG
SYLELABKLAGSGVKVKFRADQKAFQAQQLQLNFPFTILFFPKHSSKPIKYPSEK
RDVDSLMAFVNALR"
```

## transit\_peptide

106..282

/gene="par2neu"

1027..1030

/gene="par2neu"

/note="results in cysteine to tryptophan substitution"

/citation=[1]

/replace="tgt"

1072..1075

/gene="par2neu"

/note="results in glutamine to aspartic acid substitution"

/citation=[1]

/replace="gac"

1084..1086

/gene="par2neu"

/note="results in glutamic acid to glycine substitution"

/citation=[1]

/replace="gga"

1093..1095

/gene="par2neu"

/note="results in asparagine to isoleucine substitution"

/citation=[1]

/replace="ata"

1096..1098

/gene="par2neu"

/note="results in asparagine to isoleucine substitution"

/citation=[1]

/replace="ata"

1099..1101

/gene="par2neu"

/note="results in asparagine to tryptophan substitution"

/citation=[1]

/replace="cgg"

1102..1104

/gene="par2neu"

/note="frameshift results in glycine to tryptophan substitution"

/citation=[1]

/replace="cgg"

1105..1107

/gene="par2neu"

/note="results in asparagine to aspartic acid substitution"

/citation=[1]

/replace="gat"

## ORIGIN

## Alignment Scores:

Pred. No.:	1.98e-128	Length:	1753
Score:	1829.00	Matches:	353
Percent Similarity:	85.3%	Conservative:	48
Best Local Similarity:	75.1%	Mismatches:	59
Query Match:	74.8%	Indels:	10
DB:	15	Gaps:	6

US-10-731-525-8 (1-466) x CRU63784 (1-1753)

Qy	1	MetAlaLeuAlaPheThrSerSerleSer-----AlaProThrSerThrPhePro	17
Db	106	ATGGCTTTGGCTTTCACCTCTCTCAACTGCAATTCATGCTCTCTCTCTTCATTTGAA	165
Qy	18	SerSerGluProLysLeuProGlnleGlySerleArgleSerGluArgProileGly	37

Db	166	CAAAACAAAGCAGCAGCGCGCAGTTGGTTTCATTTTCAGCCTTTGGATCGGCCACATACT	225
Qy	38	GlyAlaValAsnPheAsnLeuSerGlnArgArgSerLeuValLysProValAsnAlaGlu	57
Db	226	ATATCACCTTTCCGGTGAATGTTCTCGTAGACGTTGGCCCTTAAGCCAAATAACCGCTGAG	285
Qy	58	ProProArgLysAspSerileValProLeuAlaAlaThrThrileValAlaSerAlaSer	77
Db	286	CCTAAAGGAATGAGTCAATTTGCCCTCCGCGCAACACCCGTTGCTCCCGAGGTTGAA	345
Qy	78	GluThrLysGlu---GluAspPheGluGlnIleAlaSerAspLeuAspAsnAlaSerPro	96
Db	346	GAGAAAGTTGATCTGGAAGACTATGAGAAATGGCTGATGAGCTTCAAAATGCCTCTCCT	405
Qy	97	LeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSer	116
Db	406	CTTGAGATCATGAGCAAGTCCCTGGCTTAAGTTTGGCAATGATATAGTATTGCTTTCAGT	465
Qy	117	GlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArgProPheArgVal	136
Db	466	GGGGCCGAGACGTTGCTTGATAGATATGCACACTTGACTGGTCGACCATTCAGAGTA	525
Qy	137	PheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAspAlaValGlu	156
Db	526	TTTAGCTTGGACACAGGAGATTGAATCCGAGACATATAAATTTTGTGATACGGTTGAG	585
Qy	157	LysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaValGluAlaLeu	176
Db	586	AAACAGTATGGAATTCATATTGATGATGTTTCTGATGTCGAAAGTTCAGGCTTTA	645
Qy	177	ValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGluCysCysArgVal	196
Db	646	GTTAGGAGCAGGGGCTTTTCTTCTACGAGATGCCACCAAGAGTGCTGCGGGTG	705
Qy	197	ArgLysValArgProLeuArgAlaLeuLysGlyLeuArgAlaThrIleThrGlyGln	216
Db	706	AGGAAGTTCAGGCGCTCGAAGAGCCCTTAAGGATTACGCGATGGATCACTGGGCGAG	765
Qy	217	ArgLysAspGlnSerProGlyThrArgSerGluIleProValValGlnValAspProAla	236
Db	766	AGGAAGATCAGTCTCTGGCACCTAGGTCTGAAATTCAGTTGTCAGGTTGACCCCTGT	825
Qy	237	PheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsnProValAlaAsnVal	256
Db	826	TTTGAGGAATGGATGTTGTTGAGGAGCTTGTGAGTGGATCTCTGTGCTTAATGTG	885
Qy	257	LysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAsnSerLeuHis	276
Db	886	GAAGGAAAGGATATCTGGAACCTTCTCCGACCATGGATGTGCGCGTGAAACACATTGCAC	945
Qy	277	AlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeuProGlyGln	296
Db	946	TCACAAGGTTATGTCTCAATTTGGCTGTGAGCGGTGACAGGCGGAGTCTTCCAGGGGAG	1005
Qy	297	HisGluArgGluGlyArgTrpTyrGluAspAlaLysAlaLysGluCysGlyLeuHis	316
Db	1006	CACGAGAGGAGGAGGAGTGTGTGGAGATGCCAAGGCAAGAGGAATGTGGGTATAC	1065
Qy	317	LysGlyAsnValLysGlnGlnLysGluAspValAsnGlyAsnGlyLeuSerGlnSer	336
Db	1066	AAAGGGAATATC-----AAGGAGGAAACTCTGAATAATAATGGGAATGGAGCTGTT	1116
Qy	337	HisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSerProAsnValValAsn	356
Db	1117	AATGGAATATGGA---AGTGATACCTATTGCTGATATCTTTGATACCAATAATGTACAGAG	1173
Qy	357	LeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArgLysGluProTrpLeu	376
Db	1174	TTGAGTAGGCTGGGATTTGAGAAATTTGCTGAAGTTGGAGGAAGAGAGAGCATGGCTA	1233
Qy	377	ValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMetGluGluSerTyrValAsp	396



Db 1234 GTGTACTTTACGCACCTTGGTCCGCTTTTTCGCCAGCAATGGAAGGATCATATCTTGAA 1293

Qy 397 LeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLysPheArgAlaAspGly 416

Db 1294 TTGGCAGAAAAGTTGGCTGTTCT--GGTGTAAGAGTTGGAAAGTTTAAGGCAGCGGT 1350

Qy 417 GluGlnLysGluPheAlaLysSerGluGlnLeuGlySerPheProThrIleLeuPhe 436

Db 1351 GATCAGAGGCATTTCACACAGCAAGATTGCAGCTGAATAGCTTCCCTACAAATCTTTTC 1410

Qy 437 PheProLysHisSerSerArgProThrIleLysSerProSerGluLysArgAspValAsp 456

Db 1411 TTCCCAAGCACTCATCTAGCCC--ATAAATATCCATCAGAGAAGGATGTAGAC 1467

Qy 457 SerLeuMetAlaPheValAsnAlaLeuArg 466

Db 1468 TCATTGATGCTTTTGTATTGCTCGCA 1497

RESULT 8

CQ805882 1398 bp DNA linear PAT 10-MAY-2004

DEFINITION Sequence 2293 from Patent WO2004035798.

ACCESSION CQ805882

VERSION CQ805882.1 GI:47111575

KEYWORDS Arabidopsis thaliana (thale cress)

SOURCE Arabidopsis thaliana

ORGANISM Arabidopsis thaliana

REFERENCE 1

AUTHORS Inze, D., de Veylder, L. and Vlieghe, K.

TITLE Identification of novel e2f target genes and use thereof

JOURNAL Patent: WO 2004035798-A 2293 29-APR-2004; CropDesign N.V. (BE)

FEATURES

source 1. 1398

location/Qualifiers

/organism="Arabidopsis thaliana"

/mol\_type="unassigned DNA"

/db\_xref="taxon:3702"

ORIGIN

Alignment Scores:

Pred. No.: 3.17e-127 Length: 1398

Score: 1811.50 Matches: 351

Percent Similarity: 84.0% Conservative: 48

Best Local Similarity: 73.9% Mismatches: 57

Query Match: 74.1% Indels: 19

DB: 6 Gaps: 7

US-10-731-525-8 (1-466) x CQ805882 (1-1398)

Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSer--- 19

Db 1 ATGGCAATGTCTGTAAATGTTCTTCTCTGCTTCCTGGATCATATAAATCTCGTTTC 60

Qy 20 -----GluProLysLeuProGlnIleGlySerIleArgIleSerGluArgPro 35

Db 61 GGTGTTTCATGGAGCCAAAGTTCCCAATTTGGTTCGTGAGTTATGGATCGTGT 120

Qy 36 IleGlyGlyAlaValAsnPheAsnLeuSer---GlnArgSerLeuValLysProVal 54

Db 121 CATGTTGCTCCTGTCTCTGAATCTATCTGGGAAGCATCATCATCTGTGTAAACCTTTA 180

Qy 55 AsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThrThrIleValAla 74

Db 181 AACGCTGAACCAAGCAAGAGATTCAATGATTCCTCTTGGCGCAACAATGGTAGCAGAA 240

Qy 75 SerAlaSerGluThrLysGlu-----GluAspPheGluGlnIleAlaSerAspLeu 91

Db 241 ATTGCAGAGAGATTGAGTGGTTGAGATTGAGATTGAGATTGAGATTGAGATTGAGATTG 300

Qy 92 AspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIle 111

Db 301 GAGATGCTTTCACCTCTTGAGATTATGGACAAAGCTCTTGAGAAATACGCGAATATC 360

Qy 112 AlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGly 131

Db 361 GCCATTGCAATTAGTGGTGCAGAGATGTTGCTCTTATTAGTACGCTCATTTGACATGGG 420

Qy 132 ArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeu 151

Db 421 AGGCCATTTAGAGTATTTAGTTTGGATACAGGAGGTTGAATCCTGAGACGTATCGGTTT 480

Qy 152 PheAspAlaValGlnLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaVal 171

Db 481 TTCGATGCGGTGGAGAGCACTATGGATTTAGGATTTAGTATATGTTTCTCGATTTCTGTT 540

Qy 172 GluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGln 191

Db 541 GAGGTTCAAGGTTTGGTAGGACAGGAGTGTCTCTTTTATGAGGATGGTCATCAG 600

Qy 192 GluCysCysArgValArgLysValArgProLeuArgAlaLeuLysGlyLeuArgAla 211

Db 601 GAGTGTTCGCGTCTTCCAAAGGTGAGACCTTTCAGGCGTCTCTCAAGGGGTTTAAAGGCT 660

Qy 212 TrpIleThrGlyClnArgLysAspGlnSerProGlyThrArgSerGluIleProValVal 231

Db 661 TGGATTACTGTCAGAGGAAGATCAATCTCCGGGACAAAGGCTCTGAGATTCGCGTTGTT 720

Qy 232 GlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsn 251

Db 721 CAGGTTGATCCCGTGTGTTGAGGTTTGGATGGTGGAGTTGGTAGTTGGTGGTAGGTAAT 780

Qy 252 ProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValPro 271

Db 781 CCGGTTTCCGAATCTTCAAGGGAATGATGTTTGGAACTTCTTGAGGACTATGGATGTTCCG 840

Qy 272 ValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgPro 291

Db 841 GTTAACACATTTGTCATCGCGCAGGATATATATCGATTGGATGTGAGCCTTCACGAAAGCG 900

Qy 292 ValLeuProGlyClnHisGluArgGluGlyArgTyrTrpTrpGluAspAlaLysAlaLys 311

Db 901 GTTTTACCAGGTCAGCAG 960

Qy 312 GluCysGlyLeuHisLysGlyAsnValLysGlnLysGluGluAspValAsnGlyAsn 331

Db 961 GAATGTGGACTTCACAAAGGGAATGTCAAA--GAAACTCCCATGATGCTTAAAGTGAAC 1017

Qy 332 GlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSer 351

Db 1018 GGGGAATCGAAATCC-----GCTGTTGCAGATATCTTTAAGAGT 1056

Qy 352 ProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArg 371

Db 1057 GAGAACTTGTGACTTTGAGCAGGAGGAGGATTTGAGAAATTTGATGAATTTGGAGAACCGT 1116

Qy 372 LysGluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMetGlu 391

Db 1117 AAAGAGCCTTGGATCGTCTGCTTTATGCTCGGTGGCCCTTTTGTCAAGCATGGAA 1176

Qy 392 GluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLys 411

Db 1177 GCATCGTATGATGAACCTGGCGGATAAATGGCTGGAAAGT---GGGATTAAAGTTTGCANA 1233

Qy 412 PheArgAlaAspGlyClnGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPhe 431

Db 1234 TTCAGAGCAGATGGTGCAGAGAGGAGTTGCTTAAGCAGAAATTCGAGCTCGGTAGCTTC 1293

Qy 432 ProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSerGlu 451

Db 1294 CCTACCATTTCTGGTTTCTTAAAGAACTCATCGAGACCG---ATCAAGTATTCGCTGAG 1350

Qy 452 LysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466



```
Db 1351 AAGAGAGATGTTGAGTCTTTGACTTCGTTCTTGAATCTTGTCCGA 1395
RESULT 9
AX412862 1398 bp DNA linear PAT 14-JUN-2002
LOCUS Sequence 626 from Patent WO0222675.
DEFINITION AX412862
ACCESSION AX412862
VERSION AX412862.1 GI:21445320
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE Glazebrook, J., Wang, X., Dangl, J.L., Eulgem, T. and Zhu, T.
AUTHORS Plant genes, the expression of which are altered by pathogen
TITLE infection
JOURNAL Patent: WO 0222675-A 626 21-MAR-2002;
Syngenta Participations AG (CH); UNIVERSITY OF NORTH CAROLINA AT
CHAPEL HILL (US); Glazebrook, Jan (US); Wang, Xun (US); Dangl,
Jeffrey L. (US); Eulgem, Thomas (US)
FEATURES
source
1. 1398
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
ORIGIN
Alignment Scores:
Pred. No.: 3.17e-127 Length: 1398
Score: 1811.50 Matches: 351
Percent Similarity: 84.0% Conservative: 48
Best Local Similarity: 73.9% Mismatches: 57
Query Match: 74.1% Indels: 19
DB: 6 Gaps: 7
US-10-731-525-8 (1-466) x AX412862 (1-1398)
Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSer--- 19
Db 1 ATGGCAATGTCTGAAATGTTTCTTCTTCTTCTGCTTCTGGGATCATATAAATCGTTTC 60
Qy 20 -----GluProLysLeuProGlnIleGlySerIleArgIleSerGluArgPro 35
Db 61 GGTGTTTCATTTGGAGCCAAAGATTTCGCAAAATGTTGTTGCTTGTGAGGTTATTCGATCGTGT 120
Qy 36 IleGlyGlyAlaValAsnPheAsnLeuSer---GlnArgArgSerLeuValLysProVal 54
Db 121 CATGTTGCTCTGTCGTCGAATCTATCTGGGAAGCGATCATCATCTGTTAAACCTTTA 180
Qy 55 AsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThrIleValAla 74
Db 181 AACCTGAACCAAGACAAAGAGATTCATGATTCCTCTTGGCGCAACAATGCTGACAGAA 240
Qy 75 SerAlaSerGluThrLysGlu-----GluAspPheGluGlnIleAlaSerAspLeu 91
Db 241 ATTGCAGAGGAAGTTGAAGTGGTTGAGATTGAGGATTTTGAAGAGCTTGTGAAGAAGTTA 300
Qy 92 AspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIle 111
Db 301 GAGAAATGCTTCACTCTTGAGATTATGCAAAAGCTCTTGAGAAATACGGGAACGATATC 360
Qy 112 AlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGly 131
Db 361 GCCATTGCAATTTAGTGGTGCAGAGATGTTGCTCTTATTAGTAGTACGCTCATTTGACTGG 420
Qy 132 ArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeu 151
Db 421 AGGCCATTGAGATTTAGTTTGAATACAGGAGGTTGAATCTTGAGACCGTATCGGTTT 480
Qy 152 PheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaVal 171
```

---

```
481 TTCGATCGCGTGGAGAACCACTATGGGATTAGGATTAGATATATGTTTCTCGATTCTGTT 540
Qy 172 GluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGln 191
Db 541 GAGGTTCAAGGTTTGGTTAGGACCAAGGATTTCTCTTTTATGAGGATGTCATCAG 600
Qy 192 GluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAla 211
Db 601 GAGTGTTCGCGTCTTCGAAAGGTGAGACCTTTGAGCGCTCTCTCAAGGGGTTTAAAGCT 660
Qy 212 TrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValVal 231
Db 661 TGGATTACTGTCAGAGAAAGATCAATCTCCGGGGCAAGGCTCTGAGATTCGCGGTTGTT 720
Qy 232 GluValAspProAlaPheGluGlyMetAspGlyIleGlySerLeuValLysTyrAsn 251
Db 721 CAGGTTGATCCGTTTGAAGGTTTGGATGGTGGAGTTGGTAGTTTGGTGAAGTGAAT 780
Qy 252 ProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValPro 271
Db 781 CCGGTTGCGAATGTTGAAGGGAATGATGTTTGGAACTTCTTGAGGACTATGGATGTTCCG 840
Qy 272 ValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgPro 291
Db 841 GTTAAACACATTTGCATCGCGCAGGATATATATCGATTGGATGTGAGCCTTCACGAAACG 900
Qy 292 ValLeuProGlyGlnHisGluArgGluGlyArgTyrTrpTrpGluAspAlaLysAlaLys 311
Db 901 GTTTTACCGGTCAGCACGAGAGAGAGAGATGTTGGTGGGAGAGATGCTTAAAGCCAG 960
Qy 312 GluCysGlyLeuHisLysGlyAsnValLysGlnLysGluAspValAsnGlyAsn 331
Db 961 GAATGTGACTTTCACAAAGGGAATGTCAAA--GAAAACTCCGATGATGCTTAAAGTGAAC 1017
Qy 332 GlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSer 351
Db 1018 GGGGAATCGAAATCC-----CCTGTTGCAGATATCTTTAAAGAT 1056
Qy 352 ProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArg 371
Db 1057 GAGAAATCTTGATCTTTGACAGCGCAGCGGATTTGAGAAATTTGATCAAGTTGGAGAACCGT 1116
Qy 372 LysGluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMetGlu 391
Db 1117 AAAGAGCCTTGGATCGTCGTCGTTTATGCTCGTGGTGGCCCTTTGTCGAAGCCATGAA 1176
Qy 392 GluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLys 411
Db 1177 GCATCGTATGATGAATCGCGGATAAATGGCTGGAAGT---GGGATTAAGGTTGCCAAA 1233
Qy 412 PheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPhe 431
Db 1234 TTCAGACAGATGGTGACCAAGAGAGTTGCTTAAGCAGGAAATTCGAGCTCGGTAGCTTC 1293
Qy 432 ProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSerGlu 451
Db 1294 CTTACCAATCTCGGTTTTCCTTGAACACTCATCGAGACCG---ATCAAGTATCCGCTCGAG 1350
Qy 452 LysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
Db 1351 AAGAGAGATGTTGAGTCTTTGACTTCGTTCTTGAATCTTGTCCGA 1395
RESULT 10
AX505803 1398 bp DNA linear PAT 27-SEP-2002
LOCUS Sequence 498 from Patent WO0216655.
DEFINITION AX505803
ACCESSION AX505803
VERSION AX505803.1 GI:23387040
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
```



## Alignment Scores:

Pred. No.: 3,17e-127 Length: 1398  
Score: 1811.50 Matches: 351  
Percent Similarity: 84.0% Conservatives: 48  
Best Local Similarity: 73.9% Mismatches: 57  
Query Match: 74.1% Indels: 19  
DB: 6 Gaps: 7

US-10-731-525-8 (1-466) x AX651565 (1-1398)

QY 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSer--- 19  
DB 1 ATGGCAATGCTGTAATGTTCTCTCTCGTCTTCGCGATCATAAATCTCGTTTC 60  
QY 20 -----GluProLysLeuProGlnIleGlySerIleArgIleSerGluArgPro 35  
DB 61 GGTGTTTCATTTGGAGCCAAAGATTTCGCAAAATTGGTTTGGATTATTTGGATCGTGT 120  
QY 36 IleGlyGlyAlaValAsnPheAsnLeuSer---GlnArgArgSerLeuValLysProVal 54  
DB 121 CATGTTGCTCTGTTGTTCTCTGAATCTATCTGGGAAGCGCATCATCTGTTAAACCTTTA 180  
QY 55 AsnAlaGluProProAlaGlyLysAspSerIleValProLeuAlaAlaThrThrIleValAla 74  
DB 181 AACGCTGAACCAAGACAAAGGATTCAATGATTCTCTTGGGCAACAATGGTAGCAGAA 240  
QY 75 SerAlaSerGluThrLysGlu-----GluAspPheGluGlnIleAlaSerAspLeu 91  
DB 241 ATTGCAGAGGAAGTTGAATGGTTGAGATTGAGGATTTTGAAGAGCTTGTCTAAGAGTTA 300  
QY 92 AspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIle 111  
DB 301 GAGATGCTTCACCTCTTGAGATTATGCAACAGCTCTTGAGAAATACGGGAACGATATC 360  
QY 112 AlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluIleValAlaLysLeuThrGly 131  
DB 361 GCCATTGCAATTAGTGGTGCAGAGAGTTGCTCTTATTGAGTAGCTCATTTGACTGGG 420  
QY 132 ArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrglnLeu 151  
DB 421 AGGCCATTAGAGTATTAGTTTGGATACAGGAGGTTGAATCTCGAGCATATCGGT 480  
QY 152 PheAspAlaValGluLysHisTyrglyIleArgIleGluIleMetPheProAspAlaVal 171  
DB 481 TTCATGCGGTGGAGAGCACTATGGGATTAGGATTGAGTATATGTTTCTCGATTCGTT 540  
QY 172 GluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrgluAspGlyHisGln 191  
DB 541 GAGGTTCAAGTTTGGTTAGGACCAAGGATTGTTCTCTTTTATGAGGATGGTCATCAG 600  
QY 192 GluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAla 211  
DB 601 GAGTGTTCGCGTGTTCGAAAGGTGAGACCTTTGAGCGGTGCTCTCAAGGGTTTAAAGCT 660  
QY 212 TrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValVal 231  
DB 661 TGGATTACTGTCAGAGAAAGATCAATCTCCGGGCAAGGTCGTGAGATTCCGGTGT 720  
QY 232 GlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsn 251  
DB 721 CAGGTTGATCCGGTGTGTAAGGTTTGGATGGTGGAGTTGCTAGTTTGGTGAAGTGAAT 780  
QY 252 ProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValPro 271  
DB 781 CCGGTTGCAATGTTGAAGGAATCATGTTTGGAACTTCTTGGAGACTATGAGATTGTCG 840  
QY 272 ValAsnSerLeuHisAlaLysGlyTyrrValSerIleGlyCysGluProCysThrArgPro 291  
DB 841 GTTAACATTTGCATCGCGCGAGGATATATATCGATTGATGTGAGCCTTGCACCAAGCG 900  
QY 292 ValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpTrpGluAspAlaLysAlaLys 311  
DB 901 GTTTTACCGGTTCAGCACGAGAGAGGAGAGTGGTGGTGGGAAGATGCTTAAAGCCAAAG 960

QY 312 GluCysGlyLeuHisLysGlyAsnValLysGlnLysGluAspValAsnGlyAsn 331  
DB 961 GAATGTGGACTTTCACAAAGGAATGTCAAA---GAAATCTCCGATGATGCTAAAGTGAAC 1017  
QY 332 GlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSer 351  
DB 1018 GGGGAATCGAATCC-----GCTGTTGCAGATATCTTTAAGAGT 1056  
QY 352 ProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArg 371  
DB 1057 GAGAAATCTTGTGACTTTGACGACGAGGAGATTGAGAAATTTGATGAAGTTGGAGAACCGT 1116  
QY 372 LysGluProTrpLeuValValLeuTyrrAlaProTrpCysProTyrrCysGlnAlaMetGlu 391  
DB 1117 AAAGAGCCTTGGATCGTCGCTTTATGCTCCGTGGTGGCCCCCTTTGTCAAGCAGTCGAA 1176  
QY 392 GluSerTyrrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLys 411  
DB 1177 GCATCGTATGATGAATCGCGGATAAATTCGCTGGAAGT---GGGATTAAGTTGCCAAA 1233  
QY 412 PheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPhe 431  
DB 1234 TTCAGAGCAGATGTCACCAAGAGGAGTTTGTCTAAGCAGGAATTCGAGCTCGTAGTTC 1293  
QY 432 ProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrrProSerGlu 451  
DB 1294 CCTACCATTCCTGTTTCCCTTAAGACTCATCGAGACCG---ATCAAGTATCCGCTGAG 1350  
QY 452 LysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466  
DB 1351 AAGAGAGATGTCGACTCTTGGCTTCTTGAATCTTGTCCGA 1395

## RESULT 12

BT002612 1398 bp mRNA linear PLN 23-DEC-2002  
LOCUS Arabidopsis thaliana At4g04610/F4H6\_13 mRNA, complete cds.

BT002612

BT002612

BT002612.1 GI:27363217

FLU CDNA.

Arabidopsis thaliana

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1398)

REFERENCE  
AUTHORS  
Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P.,  
Chan, M.M., Chang, C.H., Dale, J.M., Hayashizaki, Y., Huan, V.W.,  
Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,  
Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,  
Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,  
Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K.,  
Yu, G., Yuan, S., Shinzaki, K., Davis, R.W., Theologis, A. and  
Ecker, J.R.

Arabidopsis ORF clones

Unpublished

2 (bases 1 to 1398)

REFERENCE  
AUTHORS  
Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P.,  
Chan, M.M., Chang, C.H., Dale, J.M., Hayashizaki, Y., Huan, V.W.,  
Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,  
Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,  
Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,  
Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K.,  
Yu, G., Yuan, S., Shinzaki, K., Davis, R.W., Theologis, A. and  
Ecker, J.R.

Direct Submission

JOURNAL

Submitted (23-DEC-2002) Salk Institute Genomic Analysis Laboratory  
(STGAL) Plant Biology Laboratory, The Salk Institute for  
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,  
USA

COMMENT

The discrepancy does not affect the protein sequence.  
RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN

Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, RGC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Kim,C.J., Chen,H., Cheuk,R., Shim,D., Bower,J., Chan,M.M., Chang,C.M., Dale,J.M., Heuan,V.W., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S., Davis,R.W., Theologis,A., and Ecker,J.R.

Kim,C.J. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.

## FEATURES

Location/Qualifiers  
1..1398  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="4"  
/clone="U14453"  
/ecotype="Columbia"  
/note="This clone is in pUNI 51"  
1..1398  
/note="5-adenylyluliferate reductase"  
/codon\_start=1  
/product="At4g04610/F4H6 13"  
/protein\_id="AA011528.1"  
/db\_xref="GI:27363218"  
/translation="MAMSVNVSSTSSGIINSRFGVSLPKVQISGLRLLDRHVAP VSNLSGRSSVKPLNAEPTKDSMPIAATWVAETAESEVVEIEDPELAKLEN ASPLIEMDKALEKGNDAIAFGAEADVAIEVAHLTGRPFVFSLDGTGLNPETVRF FPAVEKHVGRJIEWPEPDSVQGLVRSKGLFPEYDGHQECRCRVRPLRLAKGL KAMITGQRKQSGTRSEIPVQVDPVFEGLDGVSLVKNPVPANVEGNDWNVLFRT MDVPTNTHAAGYISIGCEPCTKAVLPQHERBGRWMDAKKCEGLHKGYNKENS DAKVNGESKAVADIFKSNLVLTRQGIENLMKLENRKEPFWLVLYAPWPCQAME ASYDELADKLAGSGIKVAKPRADQDQKEFAKQBLQLGSFPFTILVFPKNSRPKYPSE KRVDSELSLFLNLVR"  
misc\_difference 507  
/note="compared to genome"  
/replace="g"

## ORIGIN

Alignment Scores:  
Pred. No.: 3,17e-127 Length: 1398  
Score: 1811.50 Matches: 351  
Percent Similarity: 84.0% Conservative: 48  
Best Local Similarity: 73.9% Mismatches: 57  
Query Match: 74.1% Indels: 19  
DB: 15 Gaps: 7  
US-10-731-525-8 (1-466) x BT002612 (1-1398)  
Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSer--- 19  
Db 1 ATGGCAATGCTGTAAATGTTCTCTCTCTCGTCTTCGGATCATATAAATCTCGTTTC 60  
Qy 20 -----GluProLysLeuProGlnIleGlySerIleArgIleSerGluArgPro 35  
Db 61 GGTGTTTCATTTGGAGCCAAAGTTTCGAAATTTGGTTTGGTGTATTTGGATCGTGT 120  
Qy 36 IleGlyGlyAlaValAenPheAenLeuSer---GlnArgArgSerLeuValLysProVal 54  
Db 121 CATGTTGCTCCTGTGCTCTGAATCTATCTATCTGGGAGCGATCATCTGTATAACCTTTA 180  
Qy 55 AsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThrThrIleValAla 74  
Db 181 AACCTGAACCAAGCAAGGATTCATGATCTCTTCTGGGCGNACAATGGTAGCAGAA 240  
Qy 75 SerAlaSerGluThrLysGlu-----GluAspPheGluGlnIleAlaSerAspLeu 91  
Db 75 SerAlaSerGluThrLysGlu-----GluAspPheGluGlnIleAlaSerAspLeu 91

Db 241 ATTGCAGAGGAAGTTGAAGTGGTTGAGATTGAGGATTTTGAAGAGCTTGTCTAAGAAGTTA 300  
Qy 92 AspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyValAsnAspIle 111  
Db 301 GAGATGCTTCACTCTTGTGATTTATGACAAAGCTCTTGAGAAATACGGGAACGATATC 360  
Qy 112 AlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGly 131  
Db 361 GCCATTGCAATTAGTGTGTGAGAGATGTTGCTCTTATTGAGTACGCTCATCTTACCTGGG 420  
Qy 132 ArgProPheArgValPheSerLeuAspThrGlyArgLeuAenProGluThrThrGlnLeu 151  
Db 421 AGGCCTATTAGATTTAGTTTGGATACAGGAGGTGTAATCTCTGAGACGTATCGGTTT 480  
Qy 152 PheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaVal 171  
Db 481 TTCATGCGGTGAGAGACACTATGGATTTAGGATTGAGTATATGTTTCTGATCTGTT 540  
Qy 172 GluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGln 191  
Db 541 GAGGTTCAAGGTTTGGTTAGGACCAAGGATGTTCTCTTTTATGAGGATGGTTCATCAG 600  
Qy 192 GluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAla 211  
Db 601 GAGTGTTCGCGTGTTCGAAAGGTGAGACCTTTGAGGCGTCTCTCAAGGGTTTAAAGGCT 660  
Qy 212 TrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValVal 231  
Db 661 TGGATTACTGCTCAGAGGAAGATCAATCTCCGGGACAAAGGCTCTGAGATTCGCGTTGTT 720  
Qy 232 GlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsn 251  
Db 721 CAGGTTGATCCGCGTGTGTTGAAGCTTGGATGGTGGATTTGGTGGTGAAGTGAAT 780  
Qy 252 ProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValPro 271  
Db 781 CCGGTTGCGAATGTTGAGGGAATGATGTTTGGAACTTCTTGAGGACTATGGATGTTCCG 840  
Qy 272 ValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgPro 291  
Db 841 GTTAAACATTTGATGCGGCGAGGGTATATATCATTTGGATGTGAGCCTTGACGAAAGCG 900  
Qy 292 ValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpGluAspAlaLysAlaLys 311  
Db 901 GTTTTACCGGCTCAGCAGCAGAGAGGAGATGGTGGTGGGAGAGATGCTTAAAGCAAG 960  
Qy 312 GluCysGlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluAspValAsnGlyAsn 331  
Db 961 GAATGTGGACTTCACAAAGGGAATGTCAA--GAAAACTCCGATGCTTAAAGTGAAC 1017  
Qy 332 GlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSer 351  
Db 1018 GGGGAATCGAAATCC-----GCTGTTGCGATATCTTTAAGAGT 1056  
Qy 352 ProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArg 371  
Db 1057 GAGAACTTGTGACTTTGAGCAGCAGGAGGATTTGAGAAATTTGATGAAGTTGGAGAACCGT 1116  
Qy 372 LysGluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMetGlu 391  
Db 1117 AAAGAGCCTTGGATCGTGTCTTATGCTCGGTGGTGGCCCTTTTGTGCAAGCATGGAA 1176  
Qy 392 GluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLys 411  
Db 1177 GCATCGATGATGATGCGCGGATTAATTTGGCTGGAGT---GGGATTAAAGTTGCCAA 1233  
Qy 412 PheArgAlaAspGlyGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPhe 431  
Db 1234 TTCAGAGCAGATGCTGACCAAGAGGATTTGCTTAAGCAGGAATTCGAGCTCGTAGCTTC 1293  
Qy 432 ProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSerGlu 451  
Db 1294 CCTACCATTTCTGTTTTTCCCTAAGAACTCATCAGAGCCG---ATCAAGTATTCGCTGAG 1350

```
Qy 452 LyAArgaspValaAepSerLeuMetAlaPheValaAenAlaLeuAArg 466
Db 1351 AAGAGAGATGTTGAGTCTTGAATCTTCTTGAATCTTGTCCGA 1395

RESULT 13
AF424582 1710 bp mRNA linear PLN 08-OCT-2001
LOCUS Arabidopsis thaliana AT4G04610/F4H6_13 mRNA, complete cds.
DEFINITION AF424582
ACCESSION AF424582
VERSION AF424582.1 GI:15983415
KEYWORDS FLI CDNA
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE
AUTHORS Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
Davis,R.W., Theologis,A. and Ecker,J.R.
Arabisdopsis cDNA clones
Unpublished
2 (bases 1 to 1710)
Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
Davis,R.W., Theologis,A. and Ecker,J.R.
Direct Submission
Submitted (26-SEP-2001) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : RIKEN
Arabisdopsis Full-Length cDNA) : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Cheuk,R., Chen,H.,
Kim,C.J., Koesema,E., Meyers,M.C., Shinn,P., Banh,J., Bowser,L.,
Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G.,
Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M.,
Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A.,
Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S.,
Davis,R.W., Theologis,A., and Ecker,J.R.

Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
contributed equally to this work as PIs.
Location/Qualifiers
1..1710
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="4"
/clone="RAFL09-34-D13(R14453)"
/ecotype="Columbia"
1..63
64..1461
/note="5'-adenylyl-leulfate reductase"
/codon_start=1
/product="AT4G04610/F4H6_13"

FEATURES
source
5'UTR
CDS
5'UTR
CDS
1462..1710
ORIGIN
Alignment Scores: 4.02e-127 Length: 1710
Pred. No.: 1811.50 Matches: 351
Score: 1811.50 Conservative: 48
Percent Similarity: 84.0% Mismatches: 57
Best Local Similarity: 73.9% Indels: 19
Query Match: 74.1% Gaps: 7
DB: 15
US-10-731-525-8 (1-466) x AF424582 (1-1710)
Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSer--- 19
Db 64 ATGCAATGTCGTAAATGTTCTTCTCTCTGCTTCTCTGGGATCATAACTCTCGTTTC 123
Qy 20 -----GluProLysLeuProGlnIleGlySerIleArgIleSerGluArgPro 35
Db 124 GGTGTTTCATTTGGAGCCAAAGTTTCGCAATTTGTTCTGTTAGGTTATTGGATCGTGT 183
Qy 36 IleGlyGlyAlaValAenAenAenSer---GlnArgArgSerLeuValLysProVal 54
Db 184 CATGTTCTCTGCTGTCTCTGAATCTATCTGGGAGCGCATCATCTCTGTTAAACCTTTA 243
Qy 55 AsnAlaGluProProArgLysAepSerIleValProLeuAlaAlaThrIleValAla 74
Db 244 AACCTGAACCAAGCAAGGATTCATGATTCCTCTTGGGGCAACATGGTAGCAGAA 303
Qy 75 SerAlaSerGluThrLysGlu-----GluAepPheGluGlnIleAlaSerAepLeu 91
Db 304 ATTGCAGAGGAAGTTGAAGTGTGAGATTGAGGATTTTGAAGAGCTTGTGAAGAAGTTA 363
Qy 92 AspAenAlaSerProLeuGluIleMetAepArgAlaLeuAepLysPheGlyAenAepIle 111
Db 364 GAGAATGCTTCACCTCTTGAGATTATGGACAAAGCTCTTGAGAAATACGGGAACCATATC 423
Qy 112 AlaIleAlaPheSerGlyAlaGluAepValAlaLeuIleGluTyrAlaLysLeuThrGly 131
Db 424 GCCATTGCATTTAGTGTGCAAGAGATTGCTCTTATTAGTAGTACGCTCATTTGACTGGG 483
Qy 132 ArgProPheArgValPheSerLeuAepThrGlyArgLeuAenProGluThrTyrGlnLeu 151
Db 484 AGCCATTAGAGATTATTAGTTTGATACAGGAGGTTGAATCTCTGAGACGATATCGGTT 543
Qy 152 PheAepAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAepAlaVal 171
Db 544 TTCGATGCGGTGGAGAGACACTATGGGATTAGGATTGATATGTTTCTTCCGATTCTGTT 603
Qy 172 GluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAepGlyHisGln 191
Db 604 GAGGTTTCAAGTTTGGTTAGGAGCAGGATGTTCTCTTTTATGAGGATGTTGTCATCAG 663
Qy 192 GluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAla 211
Db 664 GAGTGTTCGCGTGTTCGAAAGGCTGAGACCTTTGAGCGCTGCTCTCAAGGGTTTAAAGCT 723
Qy 212 TrpIleThrGlyGlnArgLysAepGlnSerProGlyThrArgSerGluIleProValVal 231
Db 724 TGGATTACTGTCAGAGGAAAGATCAATCTCCGGGGAACAAGGCTCGAGATTCGGTGTGT 783
Qy 232 GlnValAepProAlaPheGluGlyMetAepGlyGlyIleGlySerLeuValLysTrpAen 251
```









```
DB: 15 Gaps: 7
US-10-731-525-8 (1-466) x AY088011 (1-1714)
QY 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrThrThrPheProSerSer--- 19
Db 67 ATGGCAATGCTGTAATGTTCTTCTTCGCTCTTCGGGATCAAACTCTCGTTTC 126
QY 20 -----GluProLysLeuProGlnIleGlySerIleArgIleSerGluArgPro 35
Db 127 GGTGTTTCATTTGGAGCCAAAGTTTCGCAATTGTTGTTGTTGAGTTATTGGATCGGT 186
QY 36 IleGlyGlyAlaValAsnPheAsnLeuSer---GlnArgArgSerLeuValLysProVal 54
Db 187 CATGTTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 246
QY 55 AsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThrThrIleValAla 74
Db 247 AACCTGACCAAGCAAGACGATTCAATGATTCCTCTTCGGGCAACAATGGTAGCGAA 306
QY 75 SerAlaSerGluThrLysGlu-----GluAspPheGluGlnIleAlaSerAspLeu 91
Db 307 ATTGCAGAGGAGTTGAAGTGTGAGATTGAGGATTTGAGAGCTTGTGAAGAGTTA 366
QY 92 AspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIle 111
Db 367 GAGAAATGCTTACCTCTTGAGATTATGCAAAAGCTCTTGAGAAATACGGGAAACGATATC 426
QY 112 AlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGly 131
Db 427 GCATTCATTTAGTGTGAGAGATGTTGCTCTTATTGAGTACGCTCATTTGACTGGG 486
QY 132 ArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeu 151
Db 487 AGCCATTAGAGTATTAGTTTGGATACAGGAGGTTGATCTCGAGAGCTATCGGTTT 546
QY 152 PheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaVal 171
Db 547 TTCGATCGGTGGAGAGCACATATGGGATTAGGATTGAGTATATGTTTCTCGATTCTGTT 606
QY 172 GluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGln 191
Db 607 GAGTTTCAGGTTTGGTTAGGACAGGAGATTGTTCTCTTTTATGAGGATGTCATCAG 666
QY 192 GluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAla 211
Db 667 GAGTTGTCGCTGTCGAAAGGTGAGACCTTTGAGCGTCTCTCAAGGGTTTAAGGCT 726
QY 212 TrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValVal 231
Db 727 TGGATTACTGCTCAGAGCAAGATCAATCTCCGGGACAAGGTCAGATTCCCGTTGTT 786
QY 232 GlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsn 251
Db 787 CAGGTTGATCCGGTGTGAAAGTTTGGATGGTGGAGTGTGATGTTGGTGAAGTGGAA 846
QY 252 ProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValPro 271
Db 847 CCGGTTGCGAATGTTGAGGGAATGATGTTGGAATCTCTTGAAGCTATGGATGTTCCG 906
QY 272 ValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgPro 291
Db 907 GTTAAACACATTTGCATCGCGGAGGTATATATCGATTGGATGTGAGCTTGCACGAAAGCG 966
QY 292 ValLeuProGlyGlnHisGluArgGluGlyArgTyrTrpTrpGluAspAlaLys 311
Db 967 GTTTTACCCTGGTTCAGCACGAGAGGAGATGTTGGTGGGAGATGCTTAAAGCCAAG 1026
QY 312 GluCysGlyLeuHisLysGlyAsnValLysGlnLysGluLysGluAspValAsnGlyAsn 331
Db 1027 GATGTGGACTTCACAAAGGAATGTCAAA---GAAACTCCGATGATGCTAAAGTGAAC 1083
QY 332 GlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSer 351
```

Search completed: February 21, 2006, 00:18:49  
Job time : 7271 secs

**THIS PAGE LEFT BLANK**

Result No.	Score	Query #		Length	DB	ID	Description
		Match					
1	2445	100.0	1629	3	AAZ50066		Aaz50066 Soybean a
2	1951	79.8	1730	13	ADX13202		Adx13202 Plant fu
3	1923	78.7	1795	3	AAZ50065		Aaz50065 Soybean a
4	1811.5	74.1	1398	6	ABZ12693		Abz12693 Arabidops

XX  
XX

PA (DUPO ) DU PONT DE NEMOURS & CO B I.  
 XX Falco SC, Allen SM;  
 XX WPI; 2000-182427/16.  
 DR P-PSDB; AAZ50066.  
 XX Nucleic acid fragments encoding sulfate assimilation proteins in plants  
 PT and seeds useful as probes for isolating cDNAs and genes encoding  
 PT homologous proteins, and in producing transgenic plants.  
 XX Claim 3; Page 34-35; 47pp; English.  
 XX The present cDNA sequence encodes soybean adenosine-5'-phosphosulfate  
 CC (APS) reductase, a sulphate assimilation protein (SAP). The cDNA clone is  
 CC designated s12.pk0064.g4, isolated from soybean two week old developing  
 CC seedlings treated with 2.5ppm chlorimuron. APS reductase polypeptides are  
 CC useful for producing antibodies for screening cDNA expression libraries  
 CC to isolate full length APS reductase cDNA clones. Chimeric genes  
 CC comprising the APS reductase gene when transformed in a host cell can  
 CC alter the level of SAP expression. This can be used to increase sulphur  
 CC containing compounds in the cell, including the nutritionally critical  
 CC amino acids cysteine and methionine  
 XX SQ Sequence 1629 BP; 461 A; 309 C; 420 G; 439 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 2,16e-226 Length: 1629  
 Score: 2445.00 Matches: 466  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 100.0% Indels: 0  
 DB: 3 Gaps: 0

US-10-731-525-8 (1-466) x AAZ50066 (1-1629)

Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSerGlu 20  
 Db 32 ATGGCCCTCGCTTTCACCTTCTCAATTTCCGACCAATCCACCTTCCCATCATCGAA 91  
 Qy 21 ProLysLeuProGlnIleGlySerIleArgIleSerGluArgProIleGlyAlaVal 40  
 Db 92 CCCAACTTCGCAAAATGGGTCAATTAGGATTTCCGAGAGCCCATTTGGAGCGCGGT 151  
 Qy 41 AsnPheAsnLeuSerGlnArgSerLeuValLysProValAsnAlaGluProProArg 60  
 Db 152 AATTTCAATTTATCTCAAGACGGAGCTTGGTAAAGCCCGTTAAACGCCGAACCTCCACGC 211  
 Qy 61 LysAspSerIleValProLeuAlaAlaThrThrIleValAlaSerAlaSerGluThrLys 80  
 Db 212 AAGGATTCATGTTCTCTCGACGCAACACCATCGTGTCTTCTGCTCTGAGACGAA 271  
 Qy 81 GluGluAspPheGluGlnIleAlaSerAspLeuAspAsnAlaSerProLeuGluIleMet 100  
 Db 272 GAGGAGATTTTGAACAGATAGCCAGTCATCTCGACATGCTTCCACCTCTTGAATCATG 331  
 Qy 101 AspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSerGlyAlaGluAsp 120  
 Db 332 GATAGAGCCCTCGCAAAATTCGGCAACGACATAGCTATTGGCTTTCAGTGGTGTGCTGAGAT 391  
 Qy 121 ValAlaLeuIleGluTyrAlaLysLeuThrGlyArgProPheArgValPheSerLeuAsp 140  
 Db 392 GTTGCTTTGATGATGATGCGAAATTTGCGGGTCGACCTTTTAGGGTTTTCAGTTTGGAC 451  
 Qy 141 ThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAspAlaValGluLysHisTyrGly 160  
 Db 452 ACTGGGAGACTGAACCCAGAACTTATCACTTTTGTATGCGGTGAGAGCATTTATGGA 511  
 Qy 161 IleArgIleGluTyrMetPheProAspAlaValGluValGlnAlaLeuValArgSerLys 180  
 Db 512 ATTCGCAATTAGTACATGTTCCCTGATGCTGTGAGGTTTCAGGCATTTGGTGGAGTAG 571  
 Qy 181 GlyLeuPheSerPheTyrGluAspGlyHisGlnGluCysCysArgValArgLysValArg 200

Db 572 GGGTTATTCTCTTCTTACGAGGATGGCACCAAGAGTGTTCAGGGGTGACAAAGGTGAGG 631  
 Qy 201 ProLeuArgAlaLeuLysGlyLeuArgAlaTTPileThrGlyGlnArgLysAspGln 220  
 Db 632 CCTTTAAGGAGGGCCCTTAAAGGGTCTCAGAGCATGGATAACTGGTCAGAGAAAGACAG 691  
 Qy 221 SerProGlyThrArgSerGluIleProValValGlnValAspProAlaPheGluGlyMet 240  
 Db 692 TCACCTGTTACTAGTCTGAATACCGTGTTCAGGTTGATCCGGCTTTTGGAGGAATG 751  
 Qy 241 AspGlyGlyIleGlySerLeuValLysTTPAsnProValAlaAsnValLysGlyHisAsp 260  
 Db 752 GATGTTGGAATTTGGAAGCTTGGTGAAGTGAACCCCTGTTGCAAAATGTGAAGGGCCATGAC 811  
 Qy 261 IleTTPAsnPheLeuArgThrMetAsnValProValAsnSerLeuHisAlaLysGlyTyr 280  
 Db 812 ATATGGAACCTCTTAGGACCATGAATGTCCTGTGAATTCCTTGGCATGCAAAAGGATAT 871  
 Qy 281 ValSerIleGlyCysGluProCysThrArgProValLeuProGlyGlnHisGluArgGlu 300  
 Db 872 GTTTCCCAATGGGTGTGAGCCCTGCACTAGGCCCTGTTTACTGGGCAACATGAAGGGAA 931  
 Qy 301 GlyArgTTPTrpGluAspAlaLysAlaLysGluCysGlyLeuHisLysGlyAsnVal 320  
 Db 932 GGGAGGTGGTGGTGGGAGGATGCCAAAGCTAAGGAATGTGCTCTTCAAAAGGAATGTA 991  
 Qy 321 LysGlnLysGluGluAspValAsnGlyAsnGlyLeuSerGlnSerHisAlaAsnGly 340  
 Db 992 AAGCAGCAGAAAGAGGAGGATGTTAATGGAATGGGCTATCCCAATCCCATGCAATGGT 1051  
 Qy 341 AspAlaThrThrValProAspIlePheAsnSerProAsnValValAsnLeuSerArgThr 360  
 Db 1052 GATGCTACCACTGTGCTGACATTTTCAACGCCGCAATGTAGTTAATTGAGCAGGACT 1111  
 Qy 361 GlyIleGluAsnLeuAlaLysLeuGluAspArgLysGluProTrpLeuValLeuTyr 380  
 Db 1112 GGAATTCAGAAATTTGGCAAAATTTGGAGGACCGAAAGGAACCATGGCTTGTGTCTTTAT 1171  
 Qy 381 AlaProTrpCysProTyrCysGlnAlaMetGluLysSerTyrValAspLeuAlaAspLys 400  
 Db 1172 GCACCATGGTGGCCCTACTGCCAGGCTATGAGGAATCTTTATGTGTGACTTAGCAGACAG 1231  
 Qy 401 LeuAlaGlySerThrGlyMetLysValGlyLysPheArgAlaAspGlyGluGlnLysGlu 420  
 Db 1232 TTACAGGCTCAACAGGATGAAGTTGGAATTTAGACGAGATGGAGAACAGAAAGA 1291  
 Qy 421 PheAlaLysSerGluLeuGlnLeuGlySerPheProThrIleLeuPhePheProLysHis 440  
 Db 1292 TTTGCAAAAGAGTGAATGCAATTTGGAAAGCTTCCCTACGATATTATTTTCCCAAGCAT 1351  
 Qy 441 SerSerArgProThrIleLysTyrProSerGluLysArgAspValAspSerLeuMetAla 460  
 Db 1352 TCGTCTCGGCCCAACAATAAAGTATCCCTCGAANAAGAGATGTTGATTCTCTTGATGCA 1411  
 Qy 461 PheValAsnAlaLeuArg 466  
 Db 1412 TTTGTAATGCTTTAAGA 1429

RESULT 2  
 ADX13202  
 ID ADX13202 standard; cDNA; 1730 BP.  
 XX  
 AC ADX13202;  
 XX 21-APR-2005 (first entry)  
 DT  
 XX Plant full length insert polynucleotide seqid 7777.  
 DE  
 XX plant protectant; plant growth regulant; gene therapy; plant;  
 KW recombinant DNA construct; physical array; plant breeding marker;  
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;  
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;



Db 1228 GAGGAATCATATGTTGATTTGGCAAGAAAGTTAGCTGGTCA ---GGAGTGAAGTTGCA 1284  
 Qy 411 LysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySer 430  
 Db 1285 AAATTCAGCGGATGAGACCAAGGAATATGCAAAAGACTGAACCTGCAGTTGGGAAGC 1344  
 Qy 431 PheProThrLeuPhePheProLysHisSerSerArgProThrThrLeuLysTyrProSer 450  
 Db 1345 TTCCCCCAATACACTCTTCCCAACACACTCTTCTCAACCA---ATTAAAGTACCCCTCG 1401  
 Qy 451 GluLysArgAspValAspSerLeuMetAlaPheValAlaLeuArg 466  
 Db 1402 GAAAGAGAGATGTTGATTCATTGACGGCATTCTGTAATGCTTACGG 1449

## RESULT 3

AAZ50065  
 ID AAZ50065 standard; cDNA; 1795 BP.

AC AAZ50065;

XX 04-MAY-2000 (first entry)

XX Soybean adenosine-5'-phosphosulfate reductase cDNA clone se4.11g09.

XX Adenosine-5'-phosphosulfate reductase; APS reductase; soybean;  
 XX sulphate assimilation protein; SAP; clone se4.11g09; ss.

XX Glycine max.

XX Key Location/Qualifiers  
 XX CDS 50..1468  
 XX /\*tag= a  
 XX /product= "Soybean APS reductase"

XX WO200004161-A1.

XX 27-JAN-2000.

XX 13-JUL-1999; 99WO-US015808.

XX 14-JUL-1998; 98US-0092833P.

XX (DUPO ) DU PONT DE NEMOURS & CO E I.

XX Falco SC, Allen SM;

XX WPI; 2000-182427/16.

XX P-PSDB; AAY44744.

XX Nucleic acid fragments encoding sulfate assimilation proteins in plants  
 XX and seeds useful as probes for isolating cDNAs and genes encoding  
 XX homologous proteins, and in producing transgenic plants.

XX Claim 3; Page 32-33; 47pp; English.

XX The present cDNA sequence encodes soybean adenosine-5'-phosphosulfate  
 XX (APS) reductase, a sulphate assimilation protein (SAP). The cDNA clone is  
 XX designated se4.11g09, isolated from soybean embryo, 19 days after  
 XX flowering. APS reductase polypeptides are useful for producing antibodies  
 XX for screening cDNA expression libraries to isolate full length APS  
 XX reductase cDNA clones. Chimeric genes comprising the APS reductase gene  
 XX when transformed in a host cell can alter the level of SAP expression.  
 XX This can be used to increase sulphur containing compounds in the cell,  
 XX including the nutritionally critical amino acids cysteine and methionine

XX Sequence 1795 BP; 504 A; 351 C; 436 G; 504 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 8.3e-176 Length: 1795  
 Score: 1923.00 Matches: 378  
 Percent Similarity: 87.2% Conservative: 37  
 Best Local Similarity: 79.4% Mismatches: 45

Query Match:	78.7%	Indels:	16
DB:	3	Gaps:	7
US-10-731-525-8 (1-466) x AAZ50065 (1-1795)			
Qy 1 MetAlaLeuAlaPheThrSerSerIleSer-----AlaProThrSerThr 15			
Db 56 CTTGCGGTTTCCACTACTCTCTCTCTCAGTCGACGACGACGAGCGTCTCGAGTCT 115			
Qy 16 Phe-----ProSerSerGluProLysLeuProGlnIleGlySerIleArg 30			
Db 116 TTCTTCTCGGCGCTTGGATCTTTCATCGACGCTAAAGCTCGCAAAATTTGTTCTTTCGG 175			
Qy 31 IleSerGluArgProIleGlyAlaValAsnPheAsnLeuSerGlnArgIleSerLeu 50			
Db 176 TTCCCGGAGAGGCTCAAGTTTCGTCGGTGTGTTTAATTTAACTCAAGACGCTCTCTCG 235			
Qy 51 ValLysProValAsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThr 70			
Db 236 GTGAGGCCACTCAATGCCGAACCGCAACGGAATGATTCTGTGTTCTCTTGCAGCA--- 292			
Qy 71 ThrIleValAlaSerAlaSerGluThrLysGluGluAspPheGluGlnIleAlaSerAsp 90			
Db 293 ACTATCGTTTCTCTGAGGTTGAGAGGAGAAAGAAATTTTGAGCAATTTAGCGAAAGAC 352			
Qy 91 LeuAspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAsp 110			
Db 353 CTTGAAATTCATCTCTCTTGAGATTATGATTAAGCCCTCGAGAAATTTGGGAAACGAC 412			
Qy 111 IleAlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThr 130			
Db 413 ATCGCTATTGCCITTAGTGTGCTGAAGATGTTGTTGATTGATGACATTTGACG 472			
Qy 131 GlyArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGln 150			
Db 473 GGTGACCTTACAGAGTGTGTTAGTCTTGACACTGGGAGACTGAACCCAGAACTACAAA 532			
Qy 151 LeuPheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAla 170			
Db 533 TTTTGTGACGCTGTTGAGAGCATATTGGAATTCATATTGATGATACATGTTCCCTGATGCG 592			
Qy 171 ValGluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHis 190			
Db 593 GTTGAGGTTTCAGGCATTAGTAAGAACTAAGGGGCTCTTCTCATTTTACGAGATGGCAT 652			
Qy 191 GlnGluCysCysArgValArgLysValArgProLeuArgAlaLeuLysGlyLeuArg 210			
Db 653 CAAGAGTGCTGTAGAGTAAGAAAGGTGAGGCCCTTGAGGAGAGCCCTTAAGGGGTCTCAA 712			
Qy 211 AlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProVal 230			
Db 713 GCATGGATTACTGGACAGAGAAAGACAGCTCTCTGTTACTAGGTCTGAATCCCTATT 772			
Qy 231 ValGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValIleTrp 250			
Db 773 GTCCAGGTTGATCTGTTTTTTGAGGGACTGGATGGTGAATTTGGCAGCTGCTGAAGTGG 832			
Qy 251 AsnProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnVal 270			
Db 833 AACCCGGTGGCAATGTTAATGGTCTAGACATATGGAACTTCTCTTAGGACCATGAATGTT 892			
Qy 271 ProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArg 290			
Db 893 CCTGTAATTCATTGCAATCCCAAGGATATGTTTCGATTGGCTGTGACCATGACCAAGG 952			
Qy 291 ProValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpTrpGluAspAlaLysAla 310			
Db 953 CCGGTTTTTACCCGACCAACATGAAAGAGAAAGGAGTGGTGGTGGAGAGTGCCAAAGCC 1012			
Qy 311 LysGluCysGlyLeuHisLysGlyAsnValLysGlnLysGluLysGluAspValAsnGly 330			
Db 1013 AAGGAGTGTTGTTCTTCAAAAGGTAATTTGAAACAGGAAGATGTCGCCAGCTTAATGGA 1072			













PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136332P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 28-JUN-1999; 99US-0140981P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145132P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.

PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.

```

PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 4,82e-165 Length: 1712
Score: 1811.50 Matches: 351
Percent Similarity: 84.0% Conservative: 48
Best Local Similarity: 73.9% Mismatches: 57
Query Match: 74.1% Indels: 19
DB: 3 Gaps: 7

US-10-731-525-8 (1-466) x AAC47286 (1-1712)

QY 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSer--- 19
DB 67 ATGGCAATGCTGTAAATGTTCTTCTTCTCGTCTTCGGATCATATAAACTCTCGTTTC 126

QY 20 -----GluProLysLeuProGlnIleGlySerIleArgIleSerGluArgPro 35
DB 127 GGTGTTTCATTTGGAGCCAAAGTTTCGCAAAATGGTTTCTGTGAGGTTATTGGATCGGTT 186

QY 36 IleGlyGlyAlaValAsnPheAsnLeuSer---GlnArgArgSerLeuValLysProVal 54
DB 187 CATGTTGCTCTGCTCTGATCTATCTCTGGAGGATCAATCATCTGTTAAACCTTTA 246

QY 55 AsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThrIleValAla 74
DB 247 AACGCTGAACCAAGAGGATTCATGATTCCTCTTGGCGCAACAAATGGTAGCAGAA 306

QY 75 SerAlaSerGluThrLysGlu-----GluAspPheGluGlnIleAlaSerAspLeu 91
DB 307 ATTGCAGAGGAAGTTGAAGTGGTTGAGATTGAGGATTTTGAAGGCTTGTAGAAGTTA 366

QY 92 AspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIle 111
DB 367 GAGNATGCTTCACTCTTGAGATTATGGACAAAGCTCTTGAGAAATACGGGACGATATC 426

QY 112 AlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGly 131
DB 427 GCCATTGCATTTAGTGTGAGAGATGTTGCTCTTATTGAGTACGCTCATTTGACTGGG 486

QY 132 ArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeu 151
DB 487 AGGCCATTTTGAAGTATTAGTTTGGATACAGGGAGGTTGAATCCTTGAGACGTTATCGGTTT 546

QY 152 PheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaVal 171
DB 547 TTGATGCGGTGGAGAGCACTATGGATTAGATTAGATTAGATTATGTTCTCTGATTCGTT 606

QY 172 GluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGln 191
DB 607 GAGGTTCAAGGTTTGGTTAGGAGCAAGGATGTTCTCTTTTATGAGGATGTCATCAG 666

QY 192 GluCysCysArgValArgLysValArgProLeuArgAlaLeuLysGlyLeuArgAla 211
DB 667 GAGTGTTCGCGTGTTCGAAAGGTTGAGACCTTTGAGGCGTCTCTCAAGGGTTTTAAAGGCT 726

QY 212 TriPheThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValVal 231
DB 727 TGGATTACTGTCAGAGAAAGATCAATCTCCGGGCAAGGCTCTGAGATTCGGTTGTT 786

QY 232 GlnValAspProAlaPheGluGlyMetAspGlyIleGlySerLeuValLysTyrAsn 251
DB 787 CAGGTTGATCCGGTGTGTTGAGGTTTGGATGGTGGAGTTGTTGTTGGTGAAGTGAAT 846

QY 252 ProValAlaAsnValLysGlyHisAspIleTyrAsnPheLeuArgThrMetAsnValPro 271
DB 847 CCGGTTGCGAATGTTGAAGGGAATGATGTTTGGAACTTCTTCTGAGGACTATGGATGTTCCG 906

```

```

QY 272 ValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgPro 291
DB 907 GTTAAACACATTCGATCGCGCAGGGTATATATCGATTGGATGTGAGCCCTTCACGAAAGCG 966

QY 292 ValLeuProGlyGlnHisGluArgGluGlyArgTyrTrpGluAspAlaLysAlaLys 311
DB 967 GTTTTACCGGTCAGCACGAGAGAGGAGAGTGGTGGGAAAGTGTCTAAAGCCAAAG 1026

QY 312 GluCysGlyLeuHisLysGlyAsnValLysGlnLysGluGluAspValAsnGlyAsn 331
DB 1027 GAATGTGACTTCACAAAGGGAATGTCAA---GAAAACTCCGATGATGCTAAAGTGAAC 1083

QY 332 GlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSer 351
DB 1084 GGGGAATCGAAATCC-----GCTGTTGCAGATATCTTTAAGAGT 1122

QY 352 ProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArg 371
DB 1123 GAGAATCTTGTGACTTTGAGCAGCGCAGGGGATTTGAGAAATTTGATGAAGTTGGAGAACCGT 1182

QY 372 LysGluProThrLeuValValLeuTyrAlaProTyrCysProTyrCysGlnAlaMetGlu 391
DB 1183 AAAGAGCCTTGGATCGTGTGCTTTATGCTCCGTGGTCCCTTTTGTCAAGCCATCGAA 1242

QY 392 GluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLys 411
DB 1243 GCATCGTATGATGAACATGGCGGATAAATGGCTGGAAGT---GGGATTAGGTTGCCAAA 1299

QY 412 PheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPhe 431
DB 1300 TTCAGAGCAGATGCTGACCAAGAGGATTTGCTTAAGCAGGAATTCAGACTCGGTAGCTTC 1359

QY 432 ProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSerGlu 451
DB 1360 CCTACCATTTCTGTTTTCCTTAAGAACTCATCGAGACCG---ATCAAGTATCCGTCTGAG 1416

QY 452 LysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
DB 1417 AAGAGAGATGTTGAGTCTTTGACTTCTGTTCTTGAATCTTGTCGA 1461

RESULT 9
AAC37691
ID AAC37691 standard; DNA; 1714 BP.
XX
AC AAC37691;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 18317.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.

```

PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-01311449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135114P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 17-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142997P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.

PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144322P.  
PR 20-JUL-1999; 99US-0144484P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157533P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.





XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130851P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135112P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 21-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.

```

PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 22-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 4,96e-164 Length: 1702
Score: 1801.00 Matches: 348
Percent Similarity: 84.9% Conservative: 51
Best Local Similarity: 74.0% Mismatches: 55
Query Match: 73.7% Indels: 16
DB: 3 Gaps: 6

US-10-731-525-8 (1-466) x AAC38867 (1-1702)

QY 1 MetAlaLeuAlaPhe-----ThrSerSerileSerAlaProThrSerThrPhe 16
DB 78 ATGGCACTAGCAATCAACGTTTCTTCATCTCTCTCTTCGCGATCTCAAGCTCTAGCTTC 137
QY 17 ProSerSerGluProLysGluProGlnileGlySerileArgileSerGluArgProIle 36
DB 138 CCTTCTTCAGATCTCAAGTAACAAATCGGATCATAGGTTATTGAATCGTACCAT 197
QY 37 GlycylAlaValAenPheAenLeuSerGlnArgSerLeuValLysProValAenAla 56
DB 198 GTCTCTGCGGCTTCTCGAGTTTGTGCGGGAAGAGATCTCCGTGAAAGCTCTTAATGTG 257
QY 57 GluProProhArgLysAspSerileValProLeuAlaAlaThrThrileValAlaSerAla 76
DB 258 CAATCAATCAATCAAGGAATCCATTGTT-----GCTTCTGAGGTTACAGAGAAGCTA 308
QY 77 SerGluThrLysGluGluAspPheGluGlnileAlaSerAspLeuAspAenAlaSerPro 96
DB 309 GATGTGTGGAGTTGAGAACTTTGAGAACTTAGCAAGAGATTAGAGAAATGCTTCTCCT 368
QY 97 LeuGluileMetAspArgAlaLeuAspLysPheGlyAsnAspileAlaileAlaPheSer 116

```

```

DB 369 CTTGAGATCATGATAAAGCTCTTGAGAAGTTTGGGAATGACATTTGCTATTGCTTTTAGT 428
QY 117 GlyAlaGluAspValAlaLeuileGluTyAlaLysLeuThrGlyArgProPheAeqVal 136
DB 429 GGAGCTGAAGATGTTGCTCTGATTGAGTATGCTCAITTTAACTGGAAGACCTTTATAGGGTA 488
QY 137 PheSerLeuAspThrGlyArgLeuAenProGluThrTyArgLeuPheAspAlaValGlu 156
DB 489 TTCAGTTTGGATACAGGAGATTGAATCCAGAAACATATAGACTCTTCGATACCGTGGAG 548
QY 157 LysHisTyArgLysGlyLeuArgLysGlyLeuTyArgMetPheProAspAlaValGluValGlnAlaLeu 176
DB 549 AAACATTACGGTATTCGGATTGAGTATATGTTTCTCTGATGCTGTCAGGTTCAAGCTTTA 608
QY 177 ValArgSerLysGlyLeuPheSerPheTyArgLeuAspGlyHisGlnGluCysCysAeqVal 196
DB 609 GTTAGGAACCAAGGTTTGTCTCTTCTATGAAGATGCTCACCAGGAGTGTGCGGTATT 668
QY 197 ArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaTrpIleThrGlyGln 216
DB 669 AGAAGGTGAGACCTTTGAGGCGTGGTTGAAGGGTTTACGTGCTTGGATCACCTGGTCAG 728
QY 217 ArgLysAspGlnSerProGlyThrArgSerGluileProValValGlnValAlaAspProAla 236
DB 729 AGGAAAGATCAATCACCCGGGACAGATCGAGATTCCTGTTGTTGAGTTCAGTCCAGTG 788
QY 237 PheGluGlyMetAspGlyGlyLysGlySerLeuValLysTrpAenProValAlaAenVal 256
DB 789 TTTGAAGGATTAGATGGTGAGTTGGTAGTTTGGTGAAGTGAATCCGGTTGCTTAATGTT 848
QY 257 LysGlyHisAspileTrpAenPheLeuArgThrMetAenValProValAenSerLeuHis 276
DB 849 GAAGGGAATGATGTTTGAATTTCTTGAGGACCATGGATGTTCTCTGTCACACGCTTCAC 908
QY 277 AlaLysGlyTyArgValSerileGlyCysGluProCysThrArgProValLeuProGlyGln 296
DB 909 GCTGCGGTTATGTTTCAATCGGATGTGAGCCATGCACGAGAGCGGTTTTGCGGGTTCAG 968
QY 297 HisGluArgGluGlyArgTrpTrpGluAspAlaLysAlaLysGlyCysGlyLeuHis 316
DB 969 CATGAGAGAGAGGAGATGTTGTTGGGGAAGATGCTAAGGCTAAGAGTGTGCGCTTCAC 1028
QY 317 LysGlyAsnValLysGlnGlnLysGluGluAspValAenGlyAsnGlyLeuSerGlnSer 336
DB 1029 AAAGGGAATCAAGGAG-----AATACCAATGGAAACGCAACCGCT 1070
QY 337 HisAlaAsnGlyAspAlaThrValProAspilePheAenSerProAenValValAen 356
DB 1071 AATGTCACCGGACAGCC---ACAGTCCGGGATATCTTCAATAGCGAGATGTTGTGAAC 1127
QY 357 LeuSerArgThrGlyLeuGluAenLeuAlaLysLeuGluAspArgLysGluProTrpLeu 376
DB 1128 TTGACAGCAAGGAGATTGAGATTTGATGAAGTTAGAGAACCGGAAAGGCTTGGATTC 1187
QY 377 ValValLeuTyArgAlaProTrpCysProTyArgGlnAlaMetGluGluSerTyArgValAsp 396
DB 1188 GTTGTGCTTTATGACCGTGTGCGGCTTTGTCAGGCAATGGAAGCTTCATTTCGATGAA 1247
QY 397 LeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLysPheArgAlaAspGly 416
DB 1248 TTGCGCGATAAGCTGGGAGGAAGT---GGTGTGAAGTTCCTAAGTTTAGAGCTGATGTT 1304
QY 417 GluGlnLysGlyPheAlaLysSerGluLeuGlnLysGlySerPheProThrLysLeuPhe 436
DB 1305 GACCAAGAGGACTTCCCAAAAGGAGTTGACGCTTGGAGCTTTCACAACTACTTGTG 1364
QY 437 PheProLysHisSerArgProThrLysTyArgTrpSerGluLysArgAspValAsp 456
DB 1365 TTCCCAAGAGACTTTCAGAGCCA---ATCAGTATCCATCAGAGAGAGATGTTGAT 1421
QY 457 SerLeuMetAlaPheValAenAlaLeuArg 466

```

Db 1422 TCCTTGACATCTTCTTGAACTCTGTTCCG 1451  
RESULT 11  
AAC43415  
ID AAC43415 standard; DNA; 1377 BP.  
XX  
AC AAC43415;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 39173.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
XX protein identification; signal transduction pathway; metabolic pathway;  
XX promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-01231180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 21-MAY-1999; 99US-0135629P.  
PR 24-MAY-1999; 99US-0136021P.  
PR 25-MAY-1999; 99US-0136392P.  
PR 27-MAY-1999; 99US-0136782P.  
PR 28-MAY-1999; 99US-0137222P.  
PR 01-JUN-1999; 99US-0137528P.  
PR 03-JUN-1999; 99US-0137502P.  
PR 04-JUN-1999; 99US-0137724P.  
PR 07-JUN-1999; 99US-01380094P.  
PR 08-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 10-JUN-1999; 99US-0139119P.  
PR 14-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 16-JUN-1999; 99US-0139454P.  
PR 17-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 24-JUN-1999; 99US-0140823P.  
PR 28-JUN-1999; 99US-0140991P.  
PR 29-JUN-1999; 99US-0141287P.  
PR 30-JUN-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 27-JUL-1999; 99US-0145951P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147432P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.

```

PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 26-AUG-1999; 99US-0151085P.
PR 27-AUG-1999; 99US-0151086P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

```

## Alignment Scores:

```

Pred. No.: 8,99e-164
Score: 1797.00
Percent Similarity: 84.9%
Best Local Similarity: 73.8%
Query Match: 73.5%
DB: 3

```

```

Length: 1377
Matches: 347
Conservative: 52
Mismatch: 55
Indels: 16
Gaps: 6

```

US-10-731-525-8 (1-466) x AAC43415 (1-1377)

```

QY 1 MetAlaLeuAlaPhe-----ThrSerSerIleSerAlaProThrSerThrPhe 16
DB 1 ATGGCCTAGCAATCAACGTTCTTCTTCGCGATCTCAAGCTCTAGCTTC 60

```

```

QY 17 ProSerSerGluProLysLeuProGlnIleGlySerIleArgIleSerGluArgProIle 36
DB 61 CCTTCTTCAGATCTCAAAAGTAACAAAATCGGATCATTTAGGTTATTGAATCGTACCAAT 120
QY 37 GlyGlyAlaValAsnPheAsnLeuSerGlnArgSerLeuValIleProValAsnAla 56
DB 121 GTCTCTGCGGCTTCTCTGAGTTTGTCCGGAAGAGATCTCCCTGAAAGCTCTTAATGTG 180
QY 57 GluProProArgLysAspSerIleValProLeuAlaAlaThrThrIleValAlaSerAla 76
DB 181 CAATCAATTACAAGGAATCCATTGTT-----GCTTCTGAGGTTACAGAAAGCTA 231
QY 77 SerGluThrLysGluGluAspPheGluGlnIleAlaSerAspLeuAsnAlaSerPro 96
DB 232 GATGTGCTGGAAGTTCAAGACTTTGAGAACTAGCAAGAGATTAGAGAATGCTTCTCTCT 291
QY 97 LeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSer 116
DB 292 CTTGAGATCATGGAATAAGCTTTGAGAAGTTTGGGAATGACATTGCTATTGCTTTAGT 351
QY 117 GlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArgProPheArgVal 136
DB 352 GGAGCTGAGAGATGTTGCTCTGATGATGATGCTCATTTAACTGGAAGACCTTATAGGTA 411
QY 137 PheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAspAlaValGlu 156
DB 412 TTCAGTTTGGATACAGGAGATTGGAATCCAGAAACATATAGACTTTCGATACCGTGAG 471
QY 157 LysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaValGluValGlnAlaLeu 176
DB 472 AAACATTACGGTATTCGGATTGAGTATATGTTTCTGATGCTGTTGAGGTTCAAGCTTTA 531
QY 177 ValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGluCysCysArgVal 196
DB 532 GTTAGGAACAAGGTTTGTTCTCTTCTATGAAGATGGTCCAGGAGGTGTCGGTATT 591
QY 197 ArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaTrpIleThrGlyGln 216
DB 592 AGAAAGTTCAGACCTTTGAGGCGTGGTTGAAGGGTTACGTGCTTGGATCACTGGTCAG 651
QY 217 ArgLysAspGlnSerProGlyThrArgSerGluIleProValGlnValAspProAla 236
DB 652 AGGAAAGATCAATCACCCGGACCAAGATCGGAGATTCTGTTGTTTTCAGGTTGATCCAGTG 711
QY 237 PheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsnProValAlaAsnVal 256
DB 712 TTTGAAGGATTAGATGGTGGAGTTGGTAGTTGGTGAAGTGAATCCGGTTGCTAATGTT 771
QY 257 LysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAsnSerLeuHis 276
DB 772 GAAGGGAATGATGTTTGGAAATTTCTTGAGGACGATGGATGTTCTCTCAACAGCTTCAC 831
QY 277 AlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeuProGlyGln 296
DB 832 GCTGCGGTTATGTTTCAATCGGATGTGAGCCTATGACGAGACGCGGTTTTCGCGGTCAG 891
QY 297 HisGluArgGluGlyArgTrpTrpTrpGluAspAlaLysAlaLysGluCysGlyLeuHis 316
DB 892 CATGAGAGAGAAGGGAGATGGTGGTGGGAAGATGCTAAGGCTAAAGAGTGTGCGCTTCAC 951
QY 317 LysGlyAsnValLysGlnGlnLysGluGluAspValAsnGlyAsnGlyLeuSerGlnSer 336
DB 952 AAAGGGAATATCAAGGAG-----AATACCAATCGAAACGCAACCGCT 993
QY 337 HisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSerProAsnValValAsn 356
DB 994 AATGTCAACGGACAGCC---TCAGTCCGGGATATCTTCAATAGCAGGAATGTTGTGAAC 1050
QY 357 LeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArgLysGluProTrpLeu 376
DB 1051 TTGAGCAGCAAGGGGATTGAGAAATTTGATGAAGTTAGAGAACCGGAAGGCTTGGATC 1110
QY 377 ValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMetGluGluSerTyrValAsp 396

```

```
Db      1111 GTTGTGCTTTATGACCGTGTGCGCGTTTGTTCAGGCAATGGAAGCTTCATTTCGATGAA 1170      :
Qy      397 LeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLysPheArgAlaAspGly 416      :
Db      1171 TTGGCGGATAGCTGGGAGGAAGT--GGTGTGAAGGTTCTAAGTTTAGAGCTGATGGT 1227      :
Qy      417 GluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPheProThrIleLeuPhe 436      :
Db      1228 GACCAGAAGGACTTTCACAAAAGGAGTTGCAGTTTGGAAAGCTTTCACAAATACTTGTG 1287      :
Qy      437 PheProLysHisSerSerArgProThrIleLysTyrProSerGluLysArgAspValAsp 456      :
Db      1288 TTCCCAAGAACCTTCTCAAGACCA---ATCAAGTATCCATCAGACAGAGAGATGTTGAT 1344      :
Qy      457 SerLeuMetAlaPheValAsnAlaLeuArg 466      :
Db      1345 TCATTGACATCTTCTTGAATCTTGTTCGG 1374      :

RESULT 12
AAC32729
ID      AAC32729 standard; DNA; 1687 BP.
AC
AC      AAC32729;
XX
XX      17-OCT-2000 (first entry)
XX
XX      Arabidopsis thaliana DNA fragment SEQ ID NO: 432.
XX
XX      Hybridisation assay; genetic mapping; gene expression control;
XX      protein identification; signal transduction pathway; metabolic pathway;
XX      promoter; termination sequence; ss.
XX
XX      Arabidopsis thaliana.
XX
XX      EPI033405-A2.
XX
XX      06-SEP-2000.
XX
XX      25-FEB-2000; 2000EP-00301439.
XX
XX      25-FEB-1999; 99US-0121825P.
XX      05-MAR-1999; 99US-0123180P.
XX      09-MAR-1999; 99US-0123548P.
XX      23-MAR-1999; 99US-0125788P.
XX      25-MAR-1999; 99US-0126264P.
XX      29-MAR-1999; 99US-0126785P.
XX      01-APR-1999; 99US-0127462P.
XX      06-APR-1999; 99US-0128234P.
XX      08-APR-1999; 99US-0128714P.
XX      16-APR-1999; 99US-0129845P.
XX      19-APR-1999; 99US-0130077P.
XX      21-APR-1999; 99US-0130449P.
XX      23-APR-1999; 99US-0130510P.
XX      28-APR-1999; 99US-0130891P.
XX      30-APR-1999; 99US-0131449P.
XX      30-APR-1999; 99US-0132048P.
XX      30-APR-1999; 99US-0132407P.
XX      04-MAY-1999; 99US-0132484P.
XX      05-MAY-1999; 99US-0132485P.
XX      06-MAY-1999; 99US-0132486P.
XX      06-MAY-1999; 99US-0132487P.
XX      07-MAY-1999; 99US-0132863P.
XX      11-MAY-1999; 99US-0134256P.
XX      14-MAY-1999; 99US-0134218P.
XX      14-MAY-1999; 99US-0134219P.
XX      14-MAY-1999; 99US-0134221P.
XX      14-MAY-1999; 99US-0134370P.
XX      18-MAY-1999; 99US-0134768P.
XX      19-MAY-1999; 99US-0134941P.
XX      20-MAY-1999; 99US-0135124P.
XX      21-MAY-1999; 99US-0135353P.
XX      24-MAY-1999; 99US-0135629P.
XX      25-MAY-1999; 99US-0136392P.
XX      27-MAY-1999; 99US-0136782P.
XX      01-JUN-1999; 99US-0137222P.
XX      03-JUN-1999; 99US-0137528P.
XX      04-JUN-1999; 99US-0137502P.
XX      07-JUN-1999; 99US-0137724P.
XX      08-JUN-1999; 99US-0138094P.
XX      10-JUN-1999; 99US-0138540P.
XX      10-JUN-1999; 99US-0138847P.
XX      14-JUN-1999; 99US-0139119P.
XX      16-JUN-1999; 99US-0139452P.
XX      17-JUN-1999; 99US-0139453P.
XX      17-JUN-1999; 99US-0139454P.
XX      18-JUN-1999; 99US-0139455P.
XX      18-JUN-1999; 99US-0139456P.
XX      18-JUN-1999; 99US-0139457P.
XX      18-JUN-1999; 99US-0139458P.
XX      18-JUN-1999; 99US-0139459P.
XX      18-JUN-1999; 99US-0139460P.
XX      18-JUN-1999; 99US-0139461P.
XX      18-JUN-1999; 99US-0139462P.
XX      18-JUN-1999; 99US-0139463P.
XX      18-JUN-1999; 99US-0139750P.
XX      18-JUN-1999; 99US-0139763P.
XX      21-JUN-1999; 99US-0139817P.
XX      22-JUN-1999; 99US-0139899P.
XX      23-JUN-1999; 99US-0140353P.
XX      23-JUN-1999; 99US-0140354P.
XX      24-JUN-1999; 99US-0140695P.
XX      28-JUN-1999; 99US-0140823P.
XX      29-JUN-1999; 99US-0140991P.
XX      30-JUN-1999; 99US-0141287P.
XX      01-JUL-1999; 99US-0141842P.
XX      01-JUL-1999; 99US-0142154P.
XX      02-JUL-1999; 99US-0142055P.
XX      06-JUL-1999; 99US-0142390P.
XX      08-JUL-1999; 99US-0142803P.
XX      09-JUL-1999; 99US-0142920P.
XX      12-JUL-1999; 99US-0142977P.
XX      13-JUL-1999; 99US-0143542P.
XX      14-JUL-1999; 99US-0143624P.
XX      15-JUL-1999; 99US-0144005P.
XX      16-JUL-1999; 99US-0144085P.
XX      16-JUL-1999; 99US-0144086P.
XX      19-JUL-1999; 99US-0144325P.
XX      19-JUL-1999; 99US-0144331P.
XX      19-JUL-1999; 99US-0144332P.
XX      19-JUL-1999; 99US-0144333P.
XX      19-JUL-1999; 99US-0144334P.
XX      19-JUL-1999; 99US-0144335P.
XX      20-JUL-1999; 99US-0144332P.
XX      20-JUL-1999; 99US-0144632P.
XX      20-JUL-1999; 99US-0144884P.
XX      21-JUL-1999; 99US-0144814P.
XX      21-JUL-1999; 99US-0145086P.
XX      21-JUL-1999; 99US-0145088P.
XX      21-JUL-1999; 99US-0145085P.
XX      22-JUL-1999; 99US-0145087P.
XX      22-JUL-1999; 99US-0145089P.
XX      22-JUL-1999; 99US-0145192P.
XX      23-JUL-1999; 99US-0145145P.
XX      23-JUL-1999; 99US-0145218P.
XX      23-JUL-1999; 99US-0145224P.
XX      26-JUL-1999; 99US-0145276P.
XX      27-JUL-1999; 99US-0145913P.
XX      27-JUL-1999; 99US-0145918P.
XX      27-JUL-1999; 99US-0145919P.
XX      28-JUL-1999; 99US-0145951P.
XX      02-AUG-1999; 99US-0146386P.
XX      02-AUG-1999; 99US-0146388P.
XX      02-AUG-1999; 99US-0146389P.
```

```
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 08-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 08-OCT-1999; 99US-0157117P.
PR 08-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 23-OCT-1999; 99US-0161404P.
PR 23-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.

PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 6,36e-160 Length: 1687
Score: 1758.50 Matches: 353
Percent Similarity: 74.2% Conservative: 43
Best Local Similarity: 83.2% Mismatches: 49
Query Match: 71.9% Indels: 31
DB: 3 Gaps: 10

US-10-731-525-8 (1-466) x AAC32729 (1-1687)
Qy 1 MetAlaLeuAlaPheThrSerIleSerAlaProThr---SerThrPhe----- 16
Db 98 ATGGCTTTAGCTGTTACTTCTTCAACTGCAATCTCTGATCGACTTCTCTCGTCT 157
Qy 17 ---ProSerSerGluProLysLeuProGlnIleGlySerIleArgIleSerGluArgPro 35
Db 158 GGAGCTTCTTCTGAATCTAAAGCTCTCAATATGTTTCGATTAGGTTATCTGATCGAAC 217
Qy 36 IleGlyGlyAlaValAsnPheAsnLeuSerGlnArgSerLeuValLysProValAsn 55
Db 218 -----CAITGTCTCAGAGAGCTTACTCTATGAACCCCTTAAC 256
Qy 56 AlaGluPro-----ProArgLysAspSerIleValProLeuAlaAlaThrThrIleVal 73
Db 257 GCTGAGTCATCACATACAGACGAATCTTGGGTT--CTTCGTGCTTCACTCTAATT 313
Qy 74 AlaSerAlaSerGluThrLysGlu-----GluAspPheGluGlnIleAlaSerAsp 90
Db 314 GCTCTCGAAGTTTGAAGAGAAAGAGAGAGTTTGAAGACTTTTGAAGCACTTGTAAAG 373
Qy 91 LeuAspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAsp 110
Db 374 CTTGAAGATGCTTCTCCACTTGAATCATGATGATAAGCTCTTGAGAAATTTGAGAGCAA 433
Qy 111 IleAlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluIleAlaLysLeuThr 130
Db 434 ATCGCAATTGCTTTTAGTGAGCTGAAGATGTTGCATTTGATGAATATGCTCGTTAACT 493
Qy 131 GlyArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyGln 150
Db 494 GGAAGCCATTTTAGGGTTTTTAGTTTAGATACAGGAGATTAACCCGGAAGCTACAG 553
Qy 151 LeuPheAspAlaValGluLysHisTyGlyIleArgIleGluIleGluIleGluIleGluIle 170
Db 554 CTCTTTGACGCGATCGAGAGAGCTACGGGATTCGAATTTGAGTACATGTTTCTTGATGCA 613
Qy 171 ValGluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyGluAspGlyHis 190
Db 614 GTTGAGGTTCAAGCTTTAGTGAGAACAGGGTTTGTCTCATCTCTGAAGATGGTCAT 673
Qy 191 GlnGluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArg 210
Db 674 CAAGAGTTGCGCTGCGAGAAAGTTAGACCTTTGCGTCTGCTCTTAAGGGTCTTAAA 733
Qy 211 AlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProVal 230
Db 734 GCTTGATTATACAGGACAGAGAAAGACCAATCTCCGGGTACGAGATCTGAGATCCCTATT 793
Qy 231 ValGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrp 250
Db 794 GTTCAGGTTGATCCAGTGTTTGNAGGGTTAGATGGCGGTGTTGGAAGCTCTTGTGAGTGG 853
Qy 251 AsnProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnVal 270
Db 854 AATCCTTTGGCTAAATGTAAGAGCTGATGTGTGAACCTTTCTGAGAACTATGATGATGT 913
Qy 271 ProValAsnSerLeuHisAlaLysGlyTyValSerIleGlyCysGluProCysThrArg 290
Db 914 CCGGTGAATGCATTGTCACAAAGGATATGTCTCAATCGGGTGTGAGCCGTGTACTAGG 973
```



QY 291 ProValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpTrpGluAspAlaLysAla 310  
 DB |||||  
 DB 974 CCTGTGCTTCAGGTCAGCATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCT 1033  
 QY |||||  
 QY 311 LysGluCysGlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluAspValAsnGly 330  
 DB |||||  
 DB 1034 AAGAATGTGTCTACACAAAGGGAACATC-----AAGGAGGAAGAT----- 1075  
 QY |||||  
 QY 331 AsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrValProAspIlePheAsn 350  
 DB |||||  
 DB 1076 -----GGTGTGCGAGACTCAAGCCTGCTGCTGCAAGAGATATTGAA 1120  
 QY |||||  
 QY 351 SerProAsnValValAsnLeuSerArgThrGlyGlnGluAsnLeuAlaLysLeuGluAsp 370  
 DB |||||  
 DB 1121 AGCAACAATGTGTGTCATTCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCT 1180  
 QY |||||  
 QY 371 ArgLysGluProTrpLeuValValLeuValAlaProTrpCysProTyrCysGlnAlaMet 390  
 DB |||||  
 DB 1181 CGTAAGAGGGCGTGGCTGTCTACTTACGCTCTCTGTGGCCCTTTCTGCCAGCGCATG 1240  
 QY |||||  
 QY 391 GluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGly 410  
 DB |||||  
 DB 1241 GAAGCATCGTACGTGAATGGCTGAGAACTTCGGGGAATA-----CGAGTTAAAGTGGCG 1297  
 QY |||||  
 QY 411 LysPheArgAlaAspGlyGlnGlnLysGluPheAlaLysSerGluLeuGlnLysSer 430  
 DB |||||  
 DB 1298 AAATTCGAGCTGACGTGAGCAGAGAGAGTTGCTTAAGCAAGAGCTTCAGTTAGGAAGC 1357  
 QY |||||  
 QY 431 PheProThrIleLeuPhePheProLysHisSerArgProThrIleLysTyrProSer 450  
 DB |||||  
 DB 1358 TTCCCGAGCATCTCTCTTCGAAAGAGCTCCACGG---GCTATTAAAGTACCTTCA 1414  
 QY |||||  
 QY 451 GluLysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466  
 DB |||||  
 DB 1415 GAGCATAGAGATGTGATCTACTCATGTGCTTGTGTAATCTCTCTCG 1462  
 RESULT 13  
 AAF22305\_03  
 Continuation (4 of 11) of AAF22305 from base 300001 (Arabidopsis thaliana chromosome 4)  
 WP Sequence split into 11 fragments LOCUS AAF22305 Accession Aaf22305  
 WP Fragment Name Begin End  
 WP AAF22305\_00 1 110000  
 WP AAF22305\_01 100001 210000  
 WP AAF22305\_02 200001 310000  
 WP AAF22305\_03 300001 410000  
 WP AAF22305\_04 400001 510000  
 WP AAF22305\_05 500001 610000  
 WP AAF22305\_06 600001 710000  
 WP AAF22305\_07 700001 810000  
 WP AAF22305\_08 800001 910000  
 WP AAF22305\_09 900001 1010000  
 WP AAF22305\_10 1000001 1082138  
 Alignment Scores:  
 Pred. No.: 3,97e-156 Length: 110000  
 Score: 1745.50 Matches: 342  
 Percent Similarity: 80.5% Conservative: 43  
 Best Local Similarity: 71.5% Mismatches: 47  
 Query Match: 71.4% Indels: 47  
 DB: 3 Gaps: 7  
 US-10-731-525-8 (1-466) x AAF22305\_03 (1-110000)  
 QY 25 GlnIleGlySerIleArgIleSerGluArgProIleGlyAlaValAsnPheAsnLeu 44  
 DB |||||  
 DB 54155 CAATTTGGTTCGTTGAGGTATTGGATCGTGTTCATGTTGCTCTGCTCTGTAATCTA 54214  
 QY |||||  
 QY 45 Ser---GlnArgArgSerLeuValLysProValAsnAlaGluProProArgLysAspSer 63  
 DB |||||  
 DB 54215 TCTGGAGCGCATCATCTCTGTTAAACCTTTAAACGCTGAACCAACGACAAAGAGATTCA 54274  
 QY |||||  
 QY 64 IleValProLeuAlaAlaThrThrIleValAlaSerAlaSerGluThrLysGlu----- 81  
 DB |||||

DB 54275 ATGATTCCTCTTCGCGCAACAATGGTAGCAGAAATTCGACAGGAAGTTGAAGTGGTTGAG 54334  
 QY |||||  
 QY 82 ---GluAspPheGluGlnIleAlaSerAspLeuAspAsnAlaSerProLeuGluIleMet 100  
 DB |||||  
 DB 54335 ATTGAGGATTTTGAAGAGCTTGTAAAGACTTAGAGAAATGCTTCACTCTTGGAGATTATG 54394  
 QY |||||  
 QY 101 AspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSer----- 116  
 DB |||||  
 DB 54395 GACAAAGCTCTTGAGAAATACGGGACGATATGCCCATTCGATTTAG-CTATATATCGAT 54453  
 QY |||||  
 QY 116 ----- 116  
 DB 54454 CATGAGATTTGATATAAGTTTGTCTTTTGTAGTAGCTTTTGAAGAAAGATTGCTGATCTTTG 54513  
 QY |||||  
 QY 117 -----GlyAlaGluAspValAlaLeuIleGluTyrAlaLys 128  
 DB |||||  
 DB 54514 TTTTGGCTCTTGTGTGTTTCCAGTGGTGCAGAGATGTTGCTCTTATTGAGTAGCGTCAT 54573  
 QY |||||  
 DB 129 LeuThrGlyArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThr 148  
 DB |||||  
 DB 54574 TTGACTGGGAGGCCATTTAGAGTATTATTAGTTTGGATACAGGAGGTTGAATCCTGAGACG 54633  
 QY |||||  
 QY 149 TyrGlnLeuPheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPhePro 168  
 DB |||||  
 DB 54634 TATCGGTTTTTCGATCGGTGGGAGAACCACTATGGGATTAGGATTAGATATATGTTCTCT 54693  
 QY |||||  
 QY 169 AspAlaValGluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAsp 188  
 DB |||||  
 DB 54694 GATTCCTTGGAGTTCAAGGTTTGGTTAGGACGAGGAGTTGTTCTCTTTTATGAGGAT 54753  
 QY |||||  
 QY 189 GlyHisGlnGluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGly 208  
 DB |||||  
 DB 54754 GGTTCATCAGGAGTGTTCCTGCTTTCGAAAGGTGAGACCTTTGAGCGCTGCTCTCAAGGCT 54813  
 QY |||||  
 QY 209 LeuArgAlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIle 228  
 DB |||||  
 DB 54814 TTTAAGGCTTTGGATTACTGTGTGAGAGAAAGATCAATCTCCGGGCAACAAGTCTGAGATT 54873  
 QY |||||  
 QY 229 ProValValGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuVal 248  
 DB |||||  
 DB 54874 CCGGTTGTCAGGTTGATCCGGTGTGTTGAAGGTTTGGATGGTGGAGTTGGTGTGTTGGTG 54933  
 QY |||||  
 DB 249 LysTrpAsnProValAlaAsnValLysHisAspIleTrpAsnPheLeuArgThrMet 268  
 DB |||||  
 DB 54934 AAGTGGAAATCCGGTTGCGAATGTTGAAGGGAATGATGTTTGGAACTTCTTCAGGACTATG 54993  
 QY |||||  
 DB 269 AsnValProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCys 288  
 DB |||||  
 DB 54994 GATGTTCCGGTTAACACATTCATTCGATCGCGCAGGGTATATATCGATTGGATGTGAGCTTGC 55053  
 QY |||||  
 DB 289 ThrArgProValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpTrpGluAspAla 308  
 DB |||||  
 DB 55054 ACGAAGCGGTTTACCGGGTCCAGCACGAGAGAGAGGAGATGTTGGTGGAGAGATGCT 55113  
 QY |||||  
 QY 309 LysAlaLysGluCysGlyLeuHisLysGlyAsnValLysGlnLysGluAspVal 328  
 DB |||||  
 DB 55114 AAAGCCAGGAATGTGACTTCACAAAGGGAATGTCAA---GAAACTCCGATGATGCT 55170  
 QY |||||  
 QY 329 AsnGlyAsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrValProAspIle 348  
 DB |||||  
 DB 55171 AAAGTGAACGGGAATCGAAATCC-----GCTGTTGCGAGATATC 55209  
 QY |||||  
 QY 349 PheAsnSerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeu 368  
 DB |||||  
 DB 55210 TTTAAGAGTAGAATCTTGTGACTTTGAGCAGGCGAGGAGTTGAGAAATTTGATGAAGTTG 55269  
 QY |||||  
 DB 369 GluAspArgLysGluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCysGln 388  
 DB |||||  
 DB 55270 GAGAACCGTAAGAGCCCTTGGATCGTGTGTTTATGCTCCGTGGTGGCCCTTTGTCAA 55329  
 QY |||||  
 QY 389 AlaMetGluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLys 408  
 DB |||||  
 DB 55330 GCCATGGAGCATCGTATGATGAACCTGGCGGATAAATTTGGCTGGAAAGT---GGGATTAG 55386

QY 409 ValGlyLysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeu 428  
 DB 55387 GTTGCCAAATTCAGAGCAGATGGTACACAGAGAGTTTCTAAGCAGGAATTCGAGCTC 55446  
 QY 429 GlySerPheProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyr 448  
 DB 55447 GGTAGCTTCCCTACCAATTCCTGGTTTCCCTTAAAGAACATCATCGAGACCG--ATCAAGTAT 55503  
 QY 449 ProSerGluLysArgAspValAspSerLeuMetAlaPheValAlaLeuArg 466  
 DB 55504 CCCTCTGAGAGAGAGATGTTGAGTCTTTCGACTTCGTTCTTGAATCTTGTCGGA 55557

## RESULT 14

AB214500  
 ID AB214500 standard; DNA; 1365 BP.

AC AB214500;  
 XX 21-JAN-2003 (first entry)  
 DE Arabidopsis thaliana stress regulated gene SEQ ID NO 2305.  
 XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
 KW Arabidopsis thaliana;  
 XX Arabidopsis thaliana.

OS Arabidopsis thaliana.  
 PN WO200216655-A2.  
 XX 28-FEB-2002.

XX 24-AUG-2001; 2001WO-US026685.  
 PR 24-AUG-2000; 2000US-0227866P.  
 PR 26-JAN-2001; 2001US-0264647P.  
 PR 22-JUN-2001; 2001US-0300111P.

XX (SCRI ) SCRIPPS RES INST.  
 PA (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX Harper JF, Kreps J, Wang X, Zhu T;  
 PI WPI; 2002-304127/34.

XX Identifying a stress condition to which a plant cell has been exposed and  
 PT producing plants with increased tolerance to these abiotic stresses.

XX Claim 144; SEQ ID NO 2305; 577pp + Sequence Listing; English.

XX The invention relates to identifying a stress condition to which a plant  
 CC cell has been exposed, comprising: (a) contacting nucleic acid  
 CC representative of expressed polynucleotides in the plant cell with an  
 CC array or probes representative of the plant cell genome; and (b)  
 CC detecting a profile of expressed polynucleotides in the plant cell  
 CC characteristic of a stress response. The method is useful in the  
 CC production of transgenic plants, cells and seeds and in producing plants  
 CC with increased tolerance to abiotic stress. The present sequence is that  
 CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used  
 CC in methods of the invention. Note: The sequence data for this patent is  
 CC not represented in the printed specification but is based on sequence  
 CC information supplied to Derwent by the European Patent Office

XX Sequence 1365 BP; 367 A; 240 C; 374 G; 384 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	4,57e-158	Length:	1365
Score:	1738.00	Matches:	346
Percent Similarity:	82.5%	Conservative:	45
Best Local Similarity:	73.0%	Mismatches:	55
Query Match:	71.1%	Indels:	28
DB:	6	Gaps:	9

US-10-731-525-8 (1-466) x AB214500 (1-1365)

QY 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThr---SerThrPhe----- 16  
 DB 1 ATGGCTTTAGCTGTTACTTCTTCTCAACTGCAATCTCTGGATCGAGTTTCTCTCGTTCT 60  
 QY 17 ---ProSerSerGluProLysLeuProGlnIleGlySerIleArgIleSerGluArgPro 35  
 DB 61 GGAGCTTCTTCTGAATCTAAAGCTCTTCAATATGTTGATTTAGTTATCGATCGAAC 120  
 QY 36 IleGlyGlyAlaValAsnPheAsnLeuSerGlnArgSerLeuValLysProValAsn 55  
 DB 121 -----CAITTTCTCAGAGAGCTTACTCTATGAAACCTTTTAAAC 159  
 QY 56 AlaGluPro---ProArgLysAspSerIleValProLeuAlaAlaThrThrIleValAla 74  
 DB 160 GCTGAGTGCATTCACGAAGCAATCTTGGGTTTACTCGTCTTCAACTCTAATTTGCTCCT 219  
 QY 75 SerAlaSerGluThr-----LysGluGluAspPheGluGlnIleAlaSerAspLeuAsp 92  
 DB 220 GAAAGTTGAAGAGAAAGGAGGAGAAAGTTGAAGACTTTGAGCAACTTGTCTAAAGAGCTTGA 279  
 QY 93 AsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAla 112  
 DB 280 GATGCTTCTCCACTTGAATCATGGATTAAGCTCTTGAGAGATTCGGAGACCAATCGCA 339  
 QY 113 IleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArg 132  
 DB 340 ATTGCTTTTAGTGGAGCTGAAGATGTTGCATTCATTGAATATGCACGCTTTAATCTGGAAG 399  
 QY 133 ProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPhe 152  
 DB 400 CCATTTAGGGTTTTTAGTTAGATACAGGAGAGATTAAACCTCGAAACGTACAGGCTCTTT 459  
 QY 153 AspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaValGlu 172  
 DB 460 GACGAGTCCAGAGCAGTACGGGATTCGAATTGAGTACATGTTCTCTGATGAGTTGAG 519  
 QY 173 ValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGlu 192  
 DB 520 GTTCAAGCTTTAGTGAGGAACAAGGGTTGTTCTCTCATTCATGAAGATGTCATCAAGAG 579  
 QY 193 CysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaTyr 212  
 DB 580 TGTTCGCGTGTGAGGAAAGTTAGACCTTTGCGTGTCTTTAAGGGTCTTTAAAGCTTGG 639  
 QY 213 IleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValValGln 232  
 DB 640 ATTACAGGACAGAGGAAGACCAATCTCCGGGTACGAGATCTCAGATCCCTATTGTTTCA 699  
 QY 233 ValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTyrAsnPro 252  
 DB 700 GTTGTATCCAGTGTTCGAGGGTTAGATGGCGGTGTTGGAAGTCTTGTGAAGTGAATCCT 759  
 QY 253 ValAlaAsnValLysGlyHisAspIleTyrAsnPheLeuArgThrMetAsnValProVal 272  
 DB 760 TTGGCTTAATGTTGAAGAGCTGATGTGTGGAACCTTTCTGAGAACTATGATGTTCCGGTG 819  
 QY 273 AsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProVal 292  
 DB 820 AATGCATTCACGACACAGGGTATGTCAATCGGGTGTGAGCGGTGTACTAGCCCGGTG 879  
 QY 293 LeuProGlyGlnHisGluArgGluGlyArgTyrTyrTyrTyrGluAspAlaLysAlaLysGlu 312  
 DB 880 CTTCCAGGCCCAACATCAGAGAGAGAGAGAGTGTGTGTGGGAGACTGCTAAAGCTAAAGAA 939  
 QY 313 CysGlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluAspValAsnGlyAsnGly 332  
 DB 940 TGTGGTCTACAAAGGGAACATC-----AAGGAGGAAGAT----- 975  
 QY 333 LeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSerPro 352  
 DB 976 -----GGTGCTGCAGACTCAAAGCCTGTGTGTGCAAGAGATATTGAAAGCAAC 1026



```
QY 213 IleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValValGln 232
Db |||||
QY 640 ATTACAGCAGAGGAAGACCAATCTCGGGTACGAGATCTGAGATCCCTATTGTTTCAG 699
Db |||||
QY 233 ValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTyrAsnPro 252
Db |||||
QY 700 GTTGATCCAGTGTGTTGAAGGGTTAGATGGCGGTGTTGGAAGTCTTGTGAAGTGAATCCT 759
Db |||||
QY 253 ValAlaAsnValLysGlyHisAspIleTyrAsnPheLeuArgThrMetAsnValProVal 272
Db |||||
QY 760 TTGGCTAATGTTGAGGAGCTGATGTGTGNACTTCTGAGAACTATGGATGTTCCGGTG 819
Db |||||
QY 273 AsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProVal 292
Db |||||
QY 820 AATGCATTGCACGCACAAAGGTATGTCAATCGGTGTGAGCGGTGTACTAGGCCGGTG 879
Db |||||
QY 293 LeuProGlyGlnHisGluArgGluGlyArgTyrTyrGluAspAlaLysAlaLysGlu 312
Db |||||
QY 880 CTTCCAGGCCAACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 939
Db |||||
QY 313 CysGlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluAspValAsnGlyAsnGly 332
Db |||||
QY 940 TGTGCTCTACACAAAGGGAACATC-----AAGGAGGAGAGAT----- 975
Db |||||
QY 333 LeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSerPro 352
Db |||||
QY 976 -----GGTGTGTCAGACTCAAGCCTGCTGTGTCAGAGAGATATTTGAAAGCAAC 1026
Db |||||
QY 353 AsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArgLys 372
Db |||||
QY 1027 AATGTGGTTGCAATTGAGCAAAAGGAGGGGTTGAGAACTCTTTTGAAGCTAGAGAACCCGTAA 1086
Db |||||
QY 373 GluProTyrLeuValValLeuTyrAlaProTyrCysProTyrCysGlnAlaMetGluGlu 392
Db |||||
QY 1087 GAGCGTGGTGTGTCGTACTTTACGCTCCTTGTGTCCTTCTGCCAGGCGCATGGAAGCA 1146
Db |||||
QY 393 SerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLysPhe 412
Db |||||
QY 1147 TCGTACATCGAATTGGCTGAGAAACTTTCGGGAAAAA---GGAGTTAAAGTGGCGAAATTC 1203
Db |||||
QY 413 ArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPhePro 432
Db |||||
QY 1204 CGAGCTGACGGTGAGCAAGAGAGTTTGTCTAAGCAAGAGCTTCAGTTAGGGAGCTTCCCG 1263
Db |||||
QY 433 ThrIleLeuPhePheProLysHisSerArgProThrIleLysTyrProSerGluLys 452
Db |||||
QY 1264 ACGATCTTCTCTTTCGAAAAGAGCTCCACGG---GCTATTAAAGTACCCCTTCAGAGCAT 1320
Db |||||
QY 453 ArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
Db |||||
QY 1321 AGAGATGTTGATTCACTCATGTCGTTGTGAACTTCTTCGG 1362
Db |||||
```

Search completed: February 20, 2006, 22:17:32  
Job time : 906 secs

**THIS PAGE LEFT BLANK**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 20, 2006, 22:17:46 ; Search time 5571 Seconds  
(without alignments)  
3913.621 Million cell updates/sec

Title: US-10-731-525-8

Perfect score: 2445

Sequence: 1 MALAFSSISAPSTPSSSE.....KYPSEKRDVDSLMFAFNALR 466

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Zgapop 6.0 , Zgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/abs/ABSSWEB.spool/US10731525/runat.17022006.164858.123/app\_query.fasta.1  
-DB=EST -QPMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs05h  
-USPR=US10731525 @CGN 1.1 5315 @runat.17022006.164858.123 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: \*  
1: gb\_est1: \*  
2: gb\_est2: \*  
3: gb\_est3: \*  
4: gb\_est4: \*  
5: gb\_est5: \*  
6: gb\_est6: \*  
7: gb\_est7: \*  
8: gb\_est8: \*  
9: gb\_est9: \*  
10: gb\_gss2: \*  
11: gb\_gss3: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1702	69.6	1579	4	CNSOACVY BX815501 Arabidops
2	1689	69.1	1535	4	CNSOACVY BX813913 Arabidops
3	1500	61.3	1299	10	CL972938 OsIFC023
4	1346	55.1	959	7	CK264757 EST710835
5	1279	52.3	830	6	CH339330
6	1256	51.4	924	7	CV290140 aof01-6ms
7	1241.5	50.8	1308	4	AY105476 Zea mays

8	1226.5	50.2	836	9	BH652825
9	1204	49.2	910	6	CD573900
10	1187.5	48.6	831	10	CW799574
c 11	1172	47.9	764	6	CF119175
c 12	1172	47.9	764	6	CF119189
c 13	1166.5	47.7	968	7	CK263241
c 14	1163	47.6	704	6	CB340603
15	1156	47.3	856	7	CK287268
16	1149	47.0	859	8	DN982440
17	1148.5	47.0	780	3	BM111117
18	1135	46.4	858	2	BF253965
19	1130	46.2	753	5	BQ915480
c 20	1120	45.8	756	10	DU052684
c 21	1114	45.6	741	5	BQ695975
22	1111	45.4	755	5	BQ025975
23	1100	45.0	875	6	CB292552
c 24	1100	45.0	987	8	DR932581
25	1093	44.7	829	2	BF623476
26	1084	44.3	729	6	CA808814
27	1083	44.3	864	10	CG265597
28	1083	44.3	963	7	CK266388
29	1082	44.3	612	2	BE058404
30	1080.5	44.2	753	3	BQ118692
31	1077.5	44.1	719	5	BQ994338
32	1077.5	44.1	744	7	CV031415
33	1076	44.0	728	7	CV098569
34	1075.5	44.0	778	8	DN149426
35	1074.5	43.9	653	2	BG124326
c 36	1074	43.9	716	7	CO492657
37	1073.5	43.9	689	2	BG597316
38	1067.5	43.7	775	8	DN149322
c 39	1059.5	43.3	957	9	CC611389
40	1057.5	43.3	683	6	CA924320
41	1055.5	43.2	676	6	CA932068
42	1051	43.0	648	6	CD716959
43	1050.5	43.0	681	8	DR399755
44	1048	42.9	831	8	CX053400
45	1046.5	42.8	671	7	CO105768

#### ALIGNMENTS

RESULT 1	CNSOACVY	1579 bp	mrna	linear	HTC 06-FEB-2004
CNSOACVY	Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLRUS68B02 of Adult vegetative tissue of strain col-0 of Arabidopsis thaliana (thale cress).				
LOCUS	Arabidopsis thaliana				
DEFINITION	Arabidopsis thaliana (thale cress)				
ACCESSION	BX815501				
VERSION	BX815501.1				
KEYWORDS	HTC; GSIT cDNA.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
REFERENCE	1 (bases 1 to 1579)				
AUTHORS	Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.				
TITLE	Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome				
JOURNAL	Annotation				
REFERENCE	Unpublished				
AUTHORS	2 (bases 1 to 1579)				
TITLE	Genoscope.				
JOURNAL	Direct Submission				
COMMENT	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr) - Web : www.genoscope.cns.fr				

The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.

Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.

URGV INRA : Clepet C., Caboche M.

Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.

[http://www.genoscope.cns.fr/externe/sequences/Banque\\_Projet\\_EF/Full\\_length](http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length)

<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

## FEATURES

source

Location/Qualifiers

1..1579

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/clone="GSLTSL682B02"

/tissue\_type="Adult vegetative tissue"

/ecotype="Col-0"

/plasmid="pCMVSPORT\_6"

complement(1..1579)

/gene="Atlg62180"

gene

ORIGIN

## Alignment Scores:

Pred. No.: 2.33e-167 Length: 1579  
Score: 1702.00 Matches: 339  
Percent Similarity: 81.6% Conservative: 48  
Best Local Similarity: 71.5% Mismatches: 59  
Query Match: 69.6% Indels: 28  
DB: 4 Gaps: 9

US-10-731-525-8 (1-466) x CNS0ACVY (1-1579)

Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThr---SerThrPhe----- 16  
Db 90 ATGGCTTTAGCTGTTACTTCTTCTCAACTGCAATCTCTGGATCGAGTTTCTCGTTCT 149  
Qy 17 ---ProSerSerGluProLysLeuProGlnIleGlySerIleArgIleSerGluArgPro 35  
Db 150 GGAGCTTCTTCTGAATCTAAAGCTCTTCAGATATGTTTCGATTAGTTATCTGATCGAAC 209  
Qy 36 IleGlyGlyAlaValAenPheAsnLeuSerGlnArgSerLeuValLysProValAsn 55  
Db 210 -----CAITTGCTCAGAGAGCTTACTTGGGTACTCGTCTCAACTCTAATTTGCTCCT 248  
Qy 56 AlaGluPro---ProArgLysAspSerIleValProLeuAlaAlaThrIleValAla 74  
Db 249 GCTGAGTCACATTCCAGAGGGAATCTTGGGTACTCGTCTCAACTCTAATTTGCTCCT 308  
Qy 75 SerAlaSerGluThr-----LysGluGluAspPheGluGlnIleAlaSerAspLeuAsp 92  
Db 309 GAACTTGAAGAGAAAGAGGAGGAGATTGAAGACTTTGAGCAACTTGTCTAAGGAGCTTGAA 368  
Qy 93 AsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAla 112  
Db 369 GATGCTTCTCCACTTGAGATCATGATGAAGGCTCTTGAGAGATTCCGAGACCAATTCGCA 428  
Qy 113 IleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArg 132  
Db 429 ATTGCTTTAGTGAGCTGAAGATGTTGCGTTGATTGATGTCGACGTTTAACTGGAAAG 488  
Qy 133 ProPheArgValPheSerLeuAspThrGlyArgLeuAenProGluThrTyrGlnLeuPhe 152  
Db 489 CCAITTTGGGGTTTGGGTTTAGATACAGGAGATTAAACCCCTGAGACGTACGAGCTCTTT 548  
Qy 153 AspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaValGlu 172  
Db 549 GAGCGAGTCGAGAGACGATACGGGATTCGAATTGATGATCATGTTTCTTGATGCGAGTTGAG 608  
Qy 173 ValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGlu 192  
Db 609 GTTCAAGCTTTAGTGAGGAGCAAGGGTTTGTCTCATTTCTATGAAGATGGTTCATCAAGAG 668

Qy 193 CysCysArgValArgLysValArgProLeuArgAlaLeuLysGlyLeuArgAlaTyr 212  
Db 669 TGTTCGGTGTGAGAAAGTTAGACCTTTCGTCGTCTTTCGAGGCTCTTAAAGCTTGG 728  
Qy 213 IleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValValGln 232  
Db 729 ATTACAGGACAGAGAAAGACCAATCTCCGGGTACGAGATCTGAGATCCCTATTGTTTCA 788  
Qy 233 ValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTyrAsnPro 252  
Db 789 GTTGTATCCAGTGTTCGAGGGTTAGATGCGGGTGTGGAAGTCTTGTGAAGTGAATCCT 848  
Qy 253 ValAlaAsnValLysGlyHisAspIleThrAsnPheLeuArgThrMetAsnValProVal 272  
Db 849 TTGGCTAATGTTGAGGAGCTGATGTGGAATCTTCTGAGAACTACGGATGTTCCGGTG 908  
Qy 273 AsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProVal 292  
Db 909 AATGCATTGCACGCACAAAGGTTATGTCAATCGGGTGTGAGCGGTCTACTAGCCCGTG 968  
Qy 293 LeuProGlyGlnHisGluArgGluGlyArgTyrTyrTyrGluAspAlaLysAlaLysGlu 312  
Db 969 CTTCCAGGCCCAACATGAGAGAGAGGAGTGTGTGGTGGGAAGATGCTAAAGCTAAAGAA 1028  
Qy 313 CysGlyLeuHisLysGlyAsnValLysGlnLysGluGluAspValAsnGlyAsnGly 332  
Db 1029 TGTGTTCTACAAAGGGAACATC-----AAGAGGGAAGAT----- 1064  
Qy 333 LeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSerPro 352  
Db 1065 -----GGTGTGTCAGACTCAAAGCCTGCTGTCGCAAGAGATATTGAAAGCAAC 1115  
Qy 353 AsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArgLys 372  
Db 1116 AATGTGTTGCTCATTCGCAAGAGGAGGTTGAGAATCTTTTGAAGCTAGAGAACCGTAAA 1175  
Qy 373 GluProTyrLeuValValLeuTyrAlaProTyrCysProTyrCysGlnAlaMetGluGlu 392  
Db 1176 GAGCGTGGTGTGCTGCTTTCAGCTCTTGTGTCCTTCTGCGCCCTTCTCCAGGCGATGGAAGCA 1235  
Qy 393 SerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLysPhe 412  
Db 1236 TCGTACATCAATGTTGCTGAGAACTTTCGCGGAAAA---GGAGTTAAAGTCGCGAAATTC 1292  
Qy 413 ArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuLysPhePro 432  
Db 1293 CGAGCTCAGCGTGAGCAGAGAGGATTTGCTAAGCAAGAGCTTCAGTATGAGGAGCTTCCCG 1352  
Qy 433 ThrIleLeuPhePheProLysHisSerArgProThrIleLysTyrProSerGluLys 452  
Db 1353 ACATATCTTCTATTTCCGAAAAGAGCTCCAGCG---GCTATTAGTACCTTCCAGAGCAT 1409  
Qy 453 ArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466  
Db 1410 AGAGATGTTGATTCATCATGCTGCTGTTGTGAATCTTCTTCGG 1451

## RESULT 2

CNS0ADUO

LOCUS

DEFINITION

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

CNS0ADUO 1535 bp mRNA linear HTC 06-FEB-2004  
Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
GSLT6472C07 of flowers and buds of strain col-0 of Arabidopsis  
thaliana (thale cress).

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version

Keywords

Source

Organism

Accession

Version









bacterial pathogen Xylella fastidiosa

# JOURNAL COMMENT

Unpublished (2003)  
Contact: Douglas Cook, PhD  
CAES Genome Facility  
UC Davis, Plant Pathology  
One Shields Ave, Davis, CA 95616, USA  
Tel: 530 754 6561  
Fax: 530 754 6617  
Email: drcook@ucdavis.edu  
Seq primer: ACGGTACCGACATATGCC.  
Location/Qualifiers

## FEATURES

source

```
1. .830
/organism="Vitis vinifera"
/mol_type="mRNA"
/cultivar="Cabernet Sauvignon"
/db_xref="taxon:29760"
/clone="CA23EI03IVP B07"
/sex="Hermaprodite"
/dev_stage="Mid-season leaf material, collected July 25, 2001."
/lab_host="DH5alpha"
/clone_lib="Cabernet Sauvignon Leaf - CA23EI"
/note="Organ: Leaf; Vector: pDNR; Site 1: Sfil; Site 2: Sfil; CA23EI is a cDNA library of Cabernet Sauvignon leaves. The leaves were collected on July 25, 2001, in Napa Valley, California, and represent leaves in mid-season development. These leaves were verified to be infected with the bacterial pathogen, Xylella fastidiosa, based on a diagnostic assay using PCR and Xylella-specific primer pairs. The plants were asymptomatic at the time of collection, but later developed symptoms. cDNAs were made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows:
5'-AAGCAGTGGTCCAGCGCGCATTCACGCCGGG-3' and
5'-ATTCTAGAGCCGCGCGCGCATGT(30)NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."
```

## ORIGIN

Alignment Scores:  
Pred. No.: 2.43e-123 Length: 830  
Score: 1279.00 Matches: 237  
Percent Similarity: 92.1% Conservative: 21  
Best Local Similarity: 84.6% Mismatches: 16  
Query Match: 52.3% Indels: 6  
DB: 6 Gaps: 3

US-10-731-525-8 (1-466) x CB339330 (1-830)

```
Qy 82 GluAspPheGluGlnIleAlaSerAspLeuAspPheAlaSerProLeuGluileMetAsp 101
Db 9 GAGGACTATGACCAATTGGCTATGGACCTTCGAATGCTCTCTCTTGAATATGGAT 68

Qy 102 ArgAlaLeuAspLysPheGlyAspAlaIleAlaPheSerGlyValAlaGluAspVal 121
Db 69 AAGGCCCTTGCAAAATTTGGCAATGACATTGCCATTGCTTTCAGTGGTGTGAAGATATT 128

Qy 122 AlaLeuIleGluTyrAlaLysLeuThrGlyArgPropheArgValPheSerLeuAspThr 141
Db 129 GCTTTGATTGAATATGACCGCTTAACCTGCTCGCTTAGGGTATTGAGCTGGACACT 188

Qy 142 GlyArgLeuAnpProGluThrTyrGlnLeuPheAspAlaValGluLysHisTyrGlyTle 161
Db 189 GGAAGGTTGAACCCGGAACATATCAATCTTTTGACACGGTGTGAACACATATGGCATC 248

Qy 162 ArgIleGluTyrMetPheProAspAlaValGluValGlnAlaLeuValArgSerLysGly 181
Db 249 CGCATTGATACATGTTTCCAGATGCTGTTGAGTTTCAGGGCTTGGTGAGGAGCAAGGA 308

Qy 182 LeuPheSerPheTyrGluAspGlyHisGlnGluCysGlyArgValArgLysValArgPro 201
Db 309 CTGTTCTCTCTTACGAGGATGGGCACCGAGGATGCTCCCGTGTGAGGAAGGTGAGACCC 368
```

```
Qy 202 LeuArgArgAlaLeuLysGlyLeuArgAlaTrrPileThrGlyGlnArgLysAspGlnSer 221
Db 369 TTGAGGAGGCACTGAAAGGTTTGGCTGCTGATCACAGGGCAAGAGATCACTCT 428

Qy 222 ProGlyThrArgSerGluIleProValValGlnValAspProAlaPheGluGlyMetAsp 241
Db 429 CCTGGTACCAGGGCTGAAGTTCCTGTTGTCAGGTGGATCCAGCCTTCGAGGGGTTGGAT 488

Qy 242 GlyGlyIleGlySerLeuValLysTrrAsnProValAlaAsnValLysGlyHisAspIle 261
Db 489 GGTGGGGTTGGCAGCCTGGTGAATGGAACCCAGTGGCAATGTGCAGGGCATGGCATC 548

Qy 262 TrpAsnPheLeuArgThrMetAsnValProValAsnSerLeuHisAlaLysGlyTyrVal 281
Db 549 TGGAAATTCCTCGTGTCTATGAATGTGCTGTGAATTCATTCGACATCGAAGGCTATATT 608

Qy 282 SerIleGlyCysGluProCysThrArgProValLeuProGlyGlnHisGluArgGluGly 301
Db 609 TCAATTGGGTGTGAGCCATGCATAGGCTGTACTACCTGGCCAGCATGAGAGAGAAGGA 668

Qy 302 ArgTrpTrpGluAspAlaLysAlaLysGluCysGlyLeuHisLysGlyAsnValLys 321
Db 669 AGATGGTGGTGGGAAGATGCCAGGCCAAGAGTGTGGCTTTCATAAAGGAACCTCAAG 728

Qy 322 GlnGlnLysGluGluAspValAsnGlyLeuSerGlnSerHisAlaAsnGlyAsp 341
Db 729 CAGGAA---CATGGAACAACAAATGGAAATGGG-----CATGCAATGGAACC 773

Qy 342 AlaThrThrValProAspIlePheAsnSerProAsnValValAsnLeuSerArgThrGly 361
Db 774 GCC---ACTGTTCTGATCTCTTCGACACCCAGNNNTGGTTTACCTTAAACCGAGCATGGG 830
```

## RESULT 6

CV290140/c 924 bp mRNA linear EST 23-SEP-2004  
LOCUS aof01-6ms4-d06 Aof01 Asparagus officinalis cDNA clone  
DEFINITION aof01-6ms4-d06 5', mRNA sequence.

ACCESSION CV290140.1 GI:52575140

VERSION EST.

KEYWORDS Asparagus officinalis (garden asparagus)

SOURCE Asparagus officinalis

ORGANISM Asparagus officinalis

REFERENCE 1 (bases 1 to 924)

AUTHORS Genapamphillis, C., Carlson, J., Ma, H., Soltis, D., Soltis, P.,

Oppenheimer, D., Frohlich, M., Doyle, J., Tanksley, S., Webb, M.,

Leebens-Mack, J., Landherr, L., Ilut, D. and Wall, K.

Generation of ESTs from early male inflorescences of Asparagus

officinalis

Unpublished (2004)

CONTACT: Claude dePamphilis or James Leebens-Mack

Mueller Laboratory

Penn State University

208 Mueller Laboratory, Department of Biology, ATTN Rm212, Penn

State University, University Park, PA 16802, USA

Tel: 814 863 6413

Fax: 814 865 9131

Email: cwd3@psu.edu or jhl10@psu.edu

The sequence provided is trimmed of vector and low quality regions.

Full sequence and original trace file are available from the Plant

Genome Network website (<http://pgn.cornell.edu>)

Plate: aof01-6ms4 row: d column: 06

Seq primer: M13F.

Location/Qualifiers

1. .924

/organism="Asparagus officinalis"

/mol\_type="mRNA"

/db\_xref="PGEN:aof01-6ms4-d06"

/db\_xref="taxon:4686"

/clone="aof01-6ms4-d06"

/tissue\_type="male inflorescences"



```
Query Match: 50.8% Indels: 15
DB: 4 Gaps: 4
US-10-731-525-8 (1-466) x AV105476 (1-1308)

Qy 172 GluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGln 191
Db 10 GAGGTGCAGGAGCTGGTGGCGACCAAGGGCCCTCTCTCTTCTACGAGGACGGCCACGAG 69
Qy 192 GluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAla 211
Db 70 GAGTGTCTCCGGGTGCGCAGGTGCGGCCCTGCGCAGGGCGCTCAAGGGGCTTTAGGGCA 129
Qy 212 TrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValVal 231
Db 130 TGGATCACCGGCCAGAGAAAGACAGTCCCGCCGACACGAGGACGAGCATCCCCATTTGTC 189
Qy 232 GlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsn 251
Db 190 CAGGTTGATCTCTTCGAGGCTGGATGGCGGGCGCGGTAGCTTGGTCAAGTGGAAAC 249
Qy 252 ProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValPro 271
Db 250 CCCGTGGCCACAGTCGACGCAAGGACATCTGGACTTTTCTGCGGACCATGGACGTACCT 309
Qy 272 ValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgPro 291
Db 310 GTCAACACCTTGCATCTCAGGGCTACGTGTCATCGGGTCCGAGCGCGTGCACACGAGCCC 369
Qy 292 ValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpTrpGluAspAlaLysAlaLys 311
Db 370 GTCTGCGGGGACGACGAGCGGTGAAGCGGTGTGTGGAGGAGCGCCCAAGCCCAAG 429
Qy 312 GluCysGlyLeuHisLysGlyAsnValLysGlnLysGlu-----GluAsp 327
Db 430 GAGTGGCGCTTCCACAGGGCAACATTGACAAGGACGCCCGAGCGGGCGGCCCGAGGTCC 489
Qy 328 ValAsnGlyAsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrValProAsp 347
Db 490 GCCAACGCAACGGC-----TCGGCGGGCGGCCCGGAC 522
Qy 348 IlePheAsnSerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLys 367
Db 523 ATCTTCGAGAGCCCGCGGTGGTGTCTCCCTACCGCACCGGATCGAGAACTGTGTGGC 582
Qy 368 LeuGluAspArgLysGluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCys 387
Db 583 CTGGAGAACCGCGCGAGCCGTGGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 642
Qy 388 GlnAlaMetGluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMet 407
Db 643 CAGGCAATGGAGGCTCTCTAGTGGAGCTGGCGGAGAGCTGGCGGGGTCC---GGGTG 699
Qy 408 LysValGlyLysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGln 427
Db 700 AAGGTGGCCAAAGTTCGCGCGAGCGCGAGCAGAACCGCTTCGCGAGCGCGAGCTGCAG 759
Qy 428 LeuGlySerPheProThrIleLeuPhePheProLysHisSerSerArgProThrIleLys 447
Db 760 CTGCAGAGCTTTCCACCGCTGCTCTGTTCGCGGGCGCGACCGCGAGGCC---ATCAAG 816
Qy 448 TyrProSerGluLysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
Db 817 TACCCGTGGAGAGAGGAGAGCTGCATCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 873

RESULT 8
BH652825
LOCUS BH652825 BO_2_3_KB Brassica oleracea genomic clone BOHYM08,
DEFINITION genomic survey sequence.
ACCESSION BH652825
VERSION BH652825.1 GI:18710905
KEYWORDS GSS.

SOURCE Brassica oleracea
ORGANISM Brassica oleracea
REFERENCE 1 (bases 1 to 836)
AUTHORS Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,
TITLE Uterback,T.R., Wortman,J.R., White,O.R. and Town,C.D.
JOURNAL Whole genome shotgun sequencing of Brassica oleracea and its
PUBMED application to gene discovery and annotation in Arabidopsis
COMMENT Genome Res. 15 (4), 487-495 (2005)
15805490
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
FEATURES
source
location 1..836
/mol_type="genomic DNA"
/organism="Brassica oleracea"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOHYM08"
/clone_lib="BO_2_3_KB"
/notes="Vector: pPOS1; Site_1: BstXI; 2-3 kb sheared
genomic DNA inserted into pPOS1 using BstXI linkers"
ORIGIN
Alignment Scores: 8.03e-118 Length: 836
Pred. No.: 1226.50 Matches: 228
Score: 1226.50 Conservative: 23
Percent Similarity: 87.5% Mismatches: 27
Best Local Similarity: 79.4% Indels: 10
Query Match: 99.2% Gaps: 1
DB: 9

US-10-731-525-8 (1-466) x BH652825 (1-836)
Qy 151 LeuPheAspAlaValGluLysHisTyrGlyLeuArgIleGluTyrMetPheProAspAla 170
Db 2 CTCTTCGACACCGGTTC-CAACACTACGGTATTGGAATCGAGTACATGTTTCTCTGACGCT 60
Qy 171 ValGluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHis 190
Db 61 GTCGAGGTTCAGGCTTTGGTTAGAAACAAAGGTTTGTCTCTTCTACGAAAGCGGTCTAT 120
Qy 191 GlnGluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArg 210
Db 121 CAAGAGGTGTCGCCGCATCAGAAAAGTAGAGACCGGTGAGCGGTGCTTTTAAAGGTTTACGC 180
Qy 211 AlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProVal 230
Db 181 GCTTGGATCAGCGGACAAAGAAAGATCAATCACCAGGACAGATCAGAGATCCCCGTT 240
Qy 231 ValGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrp 250
Db 241 GTTCAAGTTCGATCCAGTTCCTTTTGAAGGACTACAGCGGTGAGCTGTAGCTTCGTGAAAGTGG 300
Qy 251 AsnProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnVal 270
Db 301 AACCCGGTTGGAAACGATAGAAAGGAACGATGTTTGGAGCTTTTGGAGCATATGGATGTT 360
Qy 271 ProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArg 290
Db 361 CCCGTCAACAGCTGCAGCGTGGGGGTATGTTTCCATAGGTGCGAGCGCGTGCACGAGG 420
Qy 291 ProValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpTrpGluAspAlaLysAla 310
Db |||||
```



Db 421 CGCGTTTTCAGAGTACGACGAGAGGAGGAGGTGGTGGGAGACGCTAAGGCT 480

Qy 311 LysGluCysGlyLeuHisLysGlyValLysGlnLysGluGluAspValAsnGly 330

Db 481 AAAGAGTGTGACATTCACAAAGGGAACATCAAGAGAGTAAC----- 522

Qy 331 AsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrValProAspIlePheAsn 350

Db 523 -----GCAACGGTAAATGTCAATGGAACACATCAACGGTTCGATGATATTTTCAGA 573

Qy 351 SerProAsnValValAsnLeuSerArgThrGlyTleGluAsnLeuAlaLysLeuGluAsp 370

Db 574 AGCGAAGACGTGTGACGTGAGCGGCAAGGAGTTCGAGATCTTCGATGAGCTGGAGAAC 633

Qy 371 ArgLysGluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMet 390

Db 634 AGGAAGGAGGCTTGGATCGTTGCTTTATGCGCTTGGTGGCGTTTTCGCAAGCGATG 693

Qy 391 GluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGly 410

Db 694 GAGGCTTCGTTGATGATGTTGGCAGATAAGTTGAGGGGAGGAGATGCGGTCAAGGTGGCT 753

Qy 411 LysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySer 430

Db 754 AAGTTTAGAGCTGATGAGACCAAGAGGAGTTCCGTAAGAGTGGATGGAGCTTGGGAGC 813

Qy 431 PheProThrIleLeuPhePhe 437

Db 814 TTTTCGAGGATACCTTGTGTTTC 834

RESULT 9

CD573900 910 bp mRNA linear EST 12-JUN-2003

LOCUS UCRPT01\_02\_G12\_T3 Poncirus trifoliata CTV-challenged cDNA library -

DEFINITION UCR Poncirus trifoliata cDNA clone UCRPT01\_02\_G12, mRNA sequence.

ACCESSION CD573900

VERSION 1

KEYWORDS EST.

SOURCE Poncirus trifoliata

ORGANISM Poncirus trifoliata

REFERENCE 1 (bases 1 to 910)

AUTHORS Roose,M.L., Ye,X., Federici,C.F., Close,T.J., Fenton,R.D., Wanmaker,S., Choi,Y. and Kingan,T.

TITLE Development of EST Resources and New Genetic Markers for California Citrus - Poncirus trifoliata CTV-challenged phloem - UCR Unpublished (2003)

JOURNAL Department of Botany & Plant Sciences, University of California

COMMENT Riverside, CA, 92521-0124, USA

Tel: 9097874137

Fax: 9097874437

Email: mkeal.roose@ucr.edu

Seq primer: T3.

FEATURES

source

1..910

/organism="Poncirus trifoliata"

/mol\_type="mRNA"

/cultivar="Pomeroy OP"

/db\_xref="taxon:37690"

/clone="UCRPT01\_02\_G12"

/tissue\_type="Phloem"

/dev\_stage="10-30 cm shoots"

/lab\_host="E. coli TUC121"

/clone\_lib="Poncirus trifoliata CTV-challenged cDNA library - UCR"

/note="Vector: Lambda Uni-ZAP XR, excised phagemid; Site 1: EcoRI; Site 2: XhoI; plants were grown in the greenhouse at University of California, Riverside. The scion was a open-pollinated (very probably selfed) seedling of Poncirus trifoliata cv Pomeroy that was

selected as homozygous for the Ctv resistance gene. The rootstock was sweet orange infected with citrus tristeza virus (CTV) isolate T514 over 1 year before sampling (CTV infects sweet orange, but not genotypes carrying the Ctv resistance gene. Shoots 10-30 cm long were harvested in October 2000, and the green phloem (bark) was removed and frozen quickly in dry ice. Total RNA was extracted using Trizol reagent (Gibco). Poly(A) RNA was purified, a cDNA library was made, and 0.5 million primary lambda cDNA clones were in vivo excised to give a population of pBluescript SK(-) phagemids. All steps to this point were performed in the ML roose lab at the University of California, Riverside by X. Ye. Phagemids were plated, plasmid DNA purified, cDNA clones archived, and DNA sequences determined bi-directionally using an ABI3730 at the University of California Riverside Genomics Institute, Core Instrumentation Facility, (Choi, Kingan).

Chromatogram files were transmitted to UC Riverside (by Choi), then processed at UC Riverside (by Wanmaker) using the HarVest pipeline (<http://harvest.ucr.edu>) to remove vector and cloning oligo sequences and various contaminants, and to trim to a high quality region. Sequences that retained a phred 17 region of at least 100 bases were deposited to GenBank.

## ORIGIN

Alignment Scores:

Pred. No.: 21e-115 Length: 910

Score: 1204.00 Matches: 234

Percent Similarity: 86.9% Conservative: 25

Best Local Similarity: 78.5% Mismatches: 31

Query Match: 49.2% Indels: 8

DB: 6 Gaps: 3

US-10-731-525-8 (1-466) x CD573900 (1-910)

Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPhe----- 16

Db 23 ATGGCCCTGGCTGTAACCTTCTTCAAAATTCATTTCAAGCTCCAGCTTTTGTCTTCTAGT 82

Qy 17 ProSerSerGluProLysLeuProGlnIleGlySerIleArgIleSerGluArgProIle 36

Db 83 TCTTCAAAATGACATTAAGATTTCACAAATCGGTTCTTTCGGTGTAGTGGACCG- 136

Qy 37 GlyGlyAlaValAsnPheAsnLeuSerGlnArgArgSerLeuValLysProValAsnAla 56

Db 137 GCACAGTGGCTTCACAGACCTCTCTCAACGACGATGTTTAGTAGACCGCTTGAATGCC 196

Qy 57 GluProArgLysAspSerIleValProLeuAlaAlaThrThrIleValAlaSerAla 76

Db 197 GAGCCAAAACGAAGCGATTCCGGTGGTGCCTCTCGCAGCAACTCTCGCTACCCCTGAAGTG 256

Qy 77 SerGlu-----ThrLysGluGluAspPheGluGlnIleAlaSerAspLeuAspAla 94

Db 257 GCGGAGAAAGTAGAAGGGGAGGAGGATTTTGAGCAATTTGGCTAAGGAGCTTGAATAATGCG 316

Qy 95 SerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAla 114

Db 317 TCACCTCTTGAATCATGATGATGAGCTCTTGAANAAGTTTGAANAATGACATAGCATTTGCT 376

Qy 115 PheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArgProPhe 134

Db 377 TTCAGTGGTGTGAAGATGTTGCTTTGATCGAGTATGCCCATTTAACTGGTCGACCATTT 436

Qy 135 ArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrThrTyrGlnLeuPheAspAla 154

Db 437 AGGGTGTTTAGCTCGATATCTGGAAGGCTAAACCCAGAAACATATAGATTCTTTTGATGAA 496

Qy 155 ValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaValGluValGln 174

Db 497 GTTGAGAAACATTTATGCGATACGTTATGTAATACATGTTTCTGATGCTGTTGAGTTTCA 556

Qy 175 AlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGluCysCys 194



Db 557 GCATTAGTAGGAGCAGGGGCTCTTCTCTTCTATGAGGATGGGCACAGGAGTGCTGC 616  
Qy 195 ArgValArgLysValArgProLeuArgAlaLeuLysGlyLeuArgAlaTrpIleThr 214  
Db 617 CGTGTGAGGAAAGTGAGGCCCTTCGGAGGGCCCTCAAGGGATTCGCTGCTGATCACT 676  
Qy 215 GlyClnArgLysAspGlnSerProGlyThrArgSerGluLeuProValGlnValAsp 234  
Db 677 GGCCAAAGGAAGATCATCTCCAGGTACTAGGTCTGAAATCCCTGTAGTCCAGGTGGAT 736  
Qy 235 ProAlaPheGluGlyMetAspGlyGlyLeuValSerLeuValLysTrpAsnProValAla 254  
Db 737 CCTGTTTTGAGGGGCTGGAAGTGGGGTGGCAGCCCTGGTGAAGTGAATCCAGTGGCT 796  
Qy 255 AsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAsnSer 274  
Db 797 AATCGAAGGAAATGACATTTGGAATCTCTTCGTACTATGGGTGTGAGCCATCAATTC 856  
Qy 275 LeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProVal 292  
Db 857 TTGCACCTCAAAGGGTACATCTCATTTGGGTGTGAGCCATGCACAGGCCCGTC 910

RESULT 10  
CW799574 831 bp DNA linear GSS 23-NOV-2004  
LOCUS WiscDLox339B11 Arabidopsis thaliana T-DNA insertion flanking  
DEFINITION sequences Arabidopsis thaliana genomic, genomic survey sequence.  
ACCESSION CW799574  
VERSION CW799574.1 GI:55997402  
KEYWORDS GSS.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
REFERENCE Woody,S.T., Monson,S.S., Jester,P.J., Austin-Phillips,S.,  
AUTHORS Amasino,R.M., Sussman,M.R. and Krysan,P.J.  
TITLE A New Community Resource for Knocking-Out Small Genes and  
Tandemly-Duplicated Gene Families and for Mosaic Analysis in  
Arabidopsis  
JOURNAL Unpublished (2004)  
COMMENT Contact: Woody ST  
Biotechnology Center  
University of Wisconsin-Madison  
425 Henry Mall, Madison, WI 53706, USA  
Tel: (608) 262-4640  
Email: swoody@facstaff.wisc.edu  
Class: TAIL-PCR.

FEATURES  
source  
1. .831  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/cultivar="Col-O ecotype"  
/db\_xref="taxon:3702"  
/tissue\_type="seeds produced by primary (Basta-resistant)  
transformants"  
/clone\_lib="Arabidopsis thaliana T-DNA insertion flanking  
sequences"  
/note="Vector: pBS-Lox; Sequence generated in the course  
of an Arabidopsis T-DNA tagging program. TAIL-PCR was used  
to generate sequencing templates that represent A.t.  
genomic DNA flanking the left border of the pBS-Lox T-DNA  
insert. PCR products were sequenced directly by using the  
p745 primer 5' AACGTCGCAATGTGTTATTAAGTTGTC 3'"

ORIGIN

Alignment Scores: 9.93e-114 Length: 831  
Pred. No.: 1187.50 Matches: 225  
Score: 86.1% Conservative: 23  
Percent Similarity: 78.1% Mismatches: 28  
Best Local Similarity:

Query Match: 48.6% Indels: 14  
DB: 10 Gaps: 3  
US-10-731-525-8 (1-466) x CW799574 (1-831)  
Qy 148 ThrTyrGlnLeuPheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPhe 167  
Db 5 ACNTACAGGCTCTTTGACGCANTCGAGAGCAGTACGGGATTCGAATTAGTACATGTTT 64  
Qy 168 ProAspAlaValGluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGlu 187  
Db 65 CCTGATGCAGTTGAGGTTCAAGCTTTAGTGAGGAACAAGGGTTTGTCTCATCTATGAA 124  
Qy 188 AspGlyHisGlnGluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLys 207  
Db 125 GATGGTCATCAAGAGTGTTCCTCGTGTGAGGAAGATTAGACCTTTCGTCGTCTTAAG 184  
Qy 208 GlyLeuArgAlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGlu 227  
Db 185 GGTCTTAAAGCTTGGATTACAGCACAGAGGAAAGACCAATCTCCGGGTACGAGATCTGAG 244  
Qy 228 IleProValValGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeu 247  
Db 245 ATCCCTATTGTTCCAGTTGATCCAGTGTTTGAAGGGTTAGATGGCGGTGTGGAAGTCTT 304  
Qy 248 ValLysTrpAsnProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThr 267  
Db 305 GTGAGTGGGANTCTTTGGCTATGTTGAAGGAGCTGATGTGTGGAATCTTCTGAGAACT 364  
Qy 268 MetAsnValProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluPro 287  
Db 365 ATGGATGTTCCGGTGAATGATTCAGCCACCAAGGGGTATGTGCAATCGGGTGTGAGCCG 424  
Qy 288 CysThrArgProValLeuProGlyGlnHisGluArgGluClyArgTrpTrpTrpGluAsp 307  
Db 425 TGTACTAGGCGGTGCTTCAGGCCCAACATGAGAGAGGAAGAGGTGGTGGGAAGAT 484  
Qy 308 AlaIleValAlaLysGluCysGlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluAsp 327  
Db 485 GCTAAAGCTAAAGATGTGCTTACACAAAGGGAACATC-----AAGGAGGAAGAT 535  
Qy 328 ValAsnGlyAsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrValProAsp 347  
Db 536 -----GGTGTGCAGACTCAAAAGCCTGCTGCTGCAAGAG 571  
Qy 348 IlePheAsnSerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLys 367  
Db 572 ATATTTGAAAGCAACCAATGTGGTTGCATTGAGCAAGAGGGGGTTGAGAATCTTTGAAG 631  
Qy 368 LeuGluAspArgLysGluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCys 387  
Db 632 CTAGAGAACCGTAAAGAGCGGTGGTGGTGTACTTTACGCTCTCTGCTGCTCTTCTGC 691  
Qy 388 GlnAlaMetGluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMet 407  
Db 692 CAGCGCATGGAAGCATCGTACATCGAATTGGCTGAAAAAACTTGGCGGAAAA--GGAGTT 748  
Qy 408 LysValGlyLysPheArgAlaAspGlyClnLysGluPheAlaLysSerGluLeuGln 427  
Db 749 AAAGTGGCGAAATTCNAGCTGACGGGTGAGCANAAGGA-NTTGTCTAACCANAN-CTTCAG 806  
Qy 428 LeuGlySerPheProThrIleLeu 435  
Db 807 TTAGGGAGCTTCCGACNATCTCTN 830

RESULT 11

CF119175/c  
LOCUS CF119175 764 bp mRNA linear EST 23-JUL-2003  
DEFINITION MTU10CS.P14.D12 Aspen stem cDNA Library Populus tremuloides cDNA,  
mRNA sequence.  
ACCESSION CF119175  
VERSION CF119175.1 GI:33183854  
KEYWORDS EST.

```

SOURCE
ORGANISM
Populus tremuloides (quaking aspen)
Populus tremuloides
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE
AUTHORS
1 (bases 1 to 764)
Tsai, C.-J.
TITLE
Expressed sequence tags from Aspen
JOURNAL
COMMENT
Unpublished (2003)
Contact: Tsai C-J
Plant Biotech Research Center
Michigan Technological University, School of Forest Resources &
Environmental Science
1400 Townsend Drive, Houghton, MI 49931-1295, USA
Tel: 906 487 2914
Fax: 906 487 2915
Email: chtsai@mtu.edu.

FEATURES
source
Location/Qualifiers
1..764
/organism="Populus tremuloides"
/mol_type="mRNA"
/db_xref="taxon:3693"
/clone_lib="Aspen stem cDNA Library"
/notes="Organ: stem"

ORIGIN
Alignment Scores:
Pred. No.: 3,73e-112 Length: 764
Score: 1172.00 Matches: 213
Percent Similarity: 90.6% Conservative: 18
Best Local Similarity: 83.5% Mismatches: 22
Query Match: 47.9% Indels: 2
DB: 6 Gaps: 2

US-10-731-525-8 (1-466) x CF119175 (1-764)
QY 190 HisGlnGluCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeu 209
DB 764 CATCAGGAGTGGCTGGCTGTAGGAGGTGAGACCTTGGAGCGGCTCTGAAGGGCTT 705
QY 210 ArgAlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluLeuPro 229
DB 704 CGGGCTTGGATCATCTGCCAAAGAGGATCAATCTCCAGGACGAGGTCTGAATTTCCA 645
QY 230 ValValGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLys 249
DB 644 GTTGTTCAGGTGGACCGCTTTTGGAGGATTTGGATGGTGGGCTGGAAGCCTGTGATCAAG 585
QY 250 TrpAsnProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsn 269
DB 584 TGGATCAATGGCAATGTTAGGAGCAAGATGATGAAGTCTCTCGAACCATGGAT 525
QY 270 ValProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThr 289
DB 524 GTGCCCTGTGAACCTCATTTGCATTCAGAGGATACATCTATTGTTGGCAGCCTTGACACA 465
QY 290 ArgProValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpTrpGluAspAlaLys 309
DB 464 AGGCCAGTCTTCCGGCGCCCAACATGAGAGAGAGAGGAGGTTGGTGGGAGGATGCCACA 405
QY 310 AlaLysGluCysGlyLeuHisLysGlyAsnValLysGlnLysGluGluAspValAsn 329
DB 404 GCTAAGGAATGTGGTCTTCATTAAGGAAATCTGAAACAGGGTGTATCGACGCCCACTTAT 345
QY 330 GlyAsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePhe 349
DB 344 GCAATGGAAACGGGGCTGCCCATGCAACCGG---GCTGCCACTGTGTGTGATATTTTC 288
QY 350 AsnSerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGlu 369
DB 287 AACTCCGAGAAATTTGGTCAATTTTGGACGCGCCCGGAATCGAGAACTTATTAATAATTAGAG 228

```

---

```

QY 370 AspArgLysGluProTrpLeuValValLeuTyrAlaProTrpCysPrtCysGlnAla 389
DB 227 AACCGAAAGGAACCATGGCTTGGTGGCTCTACGCCCATGGTCCCAATTTTCCCAAGGT 168
QY 390 MetGluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysVal 409
DB 167 ATGGAAGCATCCTTATGTTGAATTTGGCGGATTAAGTTAGCAGGGAGT---GGAATAAAGGTG 111
QY 410 GlyLysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGly 429
DB 110 GGAATAATTCAGGCACATGGCATCAGAGGAATTTCTCAAAACAAGAAATTCAGACTCGGA 51
QY 430 SerPheProThrIleLeuPhePheProLysHisSerSerArgPro 444
DB 50 AGTTTTCACCAATCTCTTTTCTTCTCCCTAAACACTCATCTCGCCCA 6

RESULT 12
CF119189/c
LOCUS
DEFINITION
mRNA sequence.
ACCESSION
CF119189
VERSION
CF119189.1 GI:33183882
KEYWORDS
EST.
SOURCE
Populus tremuloides (quaking aspen)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE
AUTHORS
1 (bases 1 to 764)
Ranjan, P., Kao, Y.-Y., Harding, S.A., Jiang, H., Joshi, C.P. and
Tsai, C.-J.
TITLE
Expressed sequence tags from Aspen
JOURNAL
COMMENT
Unpublished (2003)
Contact: Tsai C-J
Plant Biotech Research Center
Michigan Technological University, School of Forest Resources &
Environmental Science
1400 Townsend Drive, Houghton, MI 49931-1295, USA
Tel: 906 487 2914
Fax: 906 487 2915
Email: chtsai@mtu.edu.

FEATURES
source
Location/Qualifiers
1..764
/organism="Populus tremuloides"
/mol_type="mRNA"
/db_xref="taxon:3693"
/clone_lib="Aspen stem cDNA Library"
/notes="Organ: stem"

ORIGIN
Alignment Scores:
Pred. No.: 3,73e-112 Length: 764
Score: 1172.00 Matches: 213
Percent Similarity: 90.6% Conservative: 18
Best Local Similarity: 83.5% Mismatches: 22
Query Match: 47.9% Indels: 2
DB: 6 Gaps: 2

US-10-731-525-8 (1-466) x CF119189 (1-764)
QY 190 HisGlnGluCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeu 209
DB 764 CATCAGGAGTGGCTGGCTGTAGGAGGTGAGACCTTGGAGCGGCTCTGAAGGGCTT 705
QY 210 ArgAlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluLeuPro 229
DB 704 CGGGCTTGGATCATCTGCCAAAGAGGATCAATCTCCAGGACGAGGTCTGAATTTCCA 645
QY 230 ValValGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLys 249
DB 644 GTTGTTCAGGTGGACCGCTTTTGGAGGATTTGGATGGTGGGCTGGAAGCCTGTGATCAAG 585
QY 250 TrpAsnProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsn 269
DB 584 TGGATCAATGGCAATGTTAGGAGCAAGATGATGAAGTCTCTCGAACCATGGAT 525
QY 270 ValProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThr 289
DB 524 GTGCCCTGTGAACCTCATTTGCATTCAGAGGATACATCTATTGTTGGCAGCCTTGACACA 465
QY 290 ArgProValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpTrpGluAspAlaLys 309
DB 464 AGGCCAGTCTTCCGGCGCCCAACATGAGAGAGAGAGGAGGTTGGTGGGAGGATGCCACA 405
QY 310 AlaLysGluCysGlyLeuHisLysGlyAsnValLysGlnLysGluGluAspValAsn 329
DB 404 GCTAAGGAATGTGGTCTTCATTAAGGAAATCTGAAACAGGGTGTATCGACGCCCACTTAT 345
QY 330 GlyAsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePhe 349
DB 344 GCAATGGAAACGGGGCTGCCCATGCAACCGG---GCTGCCACTGTGTGTGATATTTTC 288
QY 350 AsnSerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGlu 369
DB 287 AACTCCGAGAAATTTGGTCAATTTTGGACGCGCCCGGAATCGAGAACTTATTAATAATTAGAG 228

```

```

Qy 250 TrpAsnProValAlaLenValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsn 269
Db 584 TGAATCCAAATGGCAATGTAGGGGCAAGATGTATGGAGTTCCTTCGACCATGGAT 525
Qy 270 ValProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThr 289
Db 524 GTGCTGTGAACATTCATTCAAAGGGATACATCTCTATTGGTTGGAGCCCTTGCCACA 465
Qy 290 ArgProValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpTrpGluAspAlaLys 309
Db 464 AGGCCAGTTCCTCGGCCCAACATGAGAGAGAAAGGAGTGGTGGTGGAGGATGCCACA 405
Qy 310 AlalysGluCysGlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluAspValAsn 329
Db 404 GCTAAGGAATGTGCTCTTCAAGGAATCTGAACACAGGATGATGCGCCCACTTAAT 345
Qy 330 GlyAsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePhe 349
Db 344 GGCAATGMAACGGGGCTGCCCATGCNAACGGG---GCTGCCACTGTTGCTGATATTTTC 288
Qy 350 AsnSerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGlu 369
Db 287 AACTCCGAGAATTTGGTCAATTTGAGCAGCCCGGAATCGAGAACTTATTAAAAATTAGAG 228
Qy 370 AspArgLysGluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCysGlnAla 389
Db 227 AACCGAAAGGAACCATGGCTTGTGTGCTCTACGCCCCCATGGTGCCTCAAGGT 168
Qy 390 MetGluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysVal 409
Db 167 ATGGAAGCATCCTATGTGAATGGCCGATAGATTACAGGGAGT---GGAAATAAGGTG 111
Qy 410 GlyLysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGly 429
Db 110 GGAATAATTCAGGCGAGATGGCATCAGAAAGAAATTCATAAAACAAGAAATTACAGCTCGGA 51
Qy 430 SerPheProThrIleLeuPhePheProLysHisSerSerArgPro 444
Db 50 AGTTTCCAAACATCTCTTTCTTCCCTTAAACACTCATCTCGCCCA 6

RESULT 13
CK263241 968 bp mRNA linear EST 03-AUG-2004
LOCUS EST709319 potato abiotic stress cdna library Solanum tuberosum CDNA
DEFINITION clone POABL30 5' end, mRNA sequence.
ACCESSION CK263241
VERSION CK263241.1 GI:39820219
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; Solanum.
REFERENCE
AUTHORS Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
TITLE Generation of ESTs from abiotic stressed potato tissue
JOURNAL Unpublished (2003)
COMMENT Other ESTs: EST709320
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/.
Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES
Location/Qualifiers
1..968
/source
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POABL30"
/tissue_type="abiotic stress treated leaf and root tissue"

```

```

/lab_host="DH10B-TonA"
/clone_lib="potato abiotic stress cdna library"
/notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
following application of the salt stress (leaves: 2hr,
6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d).
Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
applications. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots:3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d. Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d and 4d and heat-stressed roots were harvested at 6 hr,
12 hr, 1 d, and 4d. RNA was isolated from all tissues and
equal RNA from each tissue and stress was pooled to
construct the cdna library. RNA sample."

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 2e-111 Length: 968
Score: 1166.50 Matches: 230
Percent Similarity: 84.7% Conservative: 25
Best Local Similarity: 76.4% Mismatches: 35
Query Match: 47.7% Indels: 11
DB: 7 Gaps: 4

US-10-731-525-8 (1-466) x CK263241 (1-968)

Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSer--- 19
Db 80 ATGGCTTTGGCTTTTCATCTTTCACTTTTCACTGCAATTCATGGCTCTCTCTCTTCATCTTCAT 139

Qy 20 ---GluProLysLeuProGlnIleGlySerIleArgIleSerGluArgProIleGlyGly 38
Db 140 CAACAACCCATAGTATGCCAATTTGGGTAACTTTCAGCTATTGGATCAGCA----- 190

Qy 39 AlaValAsn-----PheAsnLeuSerGlnArgArgSerLeuValLysProValAsnAla 56
Db 191 -----AACAAATTTCTTGAATTTCTCTCAGAGGCGTTGTCTGTGAAGCCATTATATGCT 244

Qy 57 GluProProArgLysAspSerIleValProLeuAlaAlaThrThrIleValAlaSerAla 76
Db 245 GAACCTTAAAGGAATGATGATCCATAGTTCTCTCAGCAGCAACCTTTGTTGCTCTCTGAGGTA 304

Qy 77 SerGlu-----ThrLysGluGluAspPheGluGlnIleAlaSerAspLeuAspAsnAla 94
Db 305 ACCGAGAAAGTAATAGACCGCAGAGGATTTTGAGAACTGGCTAAGGATCTTTGAAAATGCT 364

Qy 95 SerProLeuGluLeuMetAspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAla 114
Db 365 TCACCTCTTGAGATTATGATAATGCCCTTTGAGAAATTTGGAGATGATATTGCCATGCT 424

Qy 115 PheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArgProPhe 134
Db 425 TTCACTGGTGAGAAAGATGTGCTTTGATAGATATGCACATTTACTGTGTCACCATTT 484

Qy 135 ArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAspAla 154
Db 485 AGAGTGTTCAGCTTTGATACCGGGAGGTGAATCCAGAGACCTACCAACTTTTGTATGCT 544

Qy 155 ValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaValGluValGln 174
Db 545 GTCGAGAAGCACTATGCAATTCACATCGAATACATGTTTCTTGATTTCTGTGGAAGTTCCAG 604

```

175 AlaLeuValArgSerIysGlyLeuPheSerPheTyrGluAspGlyHisGlnGluCys 194  
 Db GCATTAGTAAGGACCAAGGGCTCTTTCTCATTTTACGAGGATGGCCACCAAGATGCTGC 664  
 195 ArgValArgIysValArgProLeuArgArgAlaLeuIysGlyLeuArgAlaTrpIleThr 214  
 Db CGTGTAGGAAGATTAGCCCTTGGAGAGCCCTCAAGAGTTGGCTGCTGGATCAGC 724  
 215 GlyGlnArgIysAspGlnSerProGlyThrArgSerGluIleProValValGlnValAsp 234  
 Db GGGCAAGAAAGATCATGCTCCCTGGAACTCGATCTGAAAGTTCAGTTGTCAGTAGAC 784  
 235 ProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValIysTrpAsnProValAla 254  
 Db CCCTCTTCGAGGATTTGGATGGTGGCTCTGGAAGCTTGGTGAAGTGGAAACCCGGTGGCT 844  
 255 AsnValIysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAsnSer 274  
 Db AATGTAGAGGCAAGGACATATGGAATCTCTAGCTGCAATGAATGTTCCCGTGAATCT 904  
 275 LeuHisAlaIysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeuPro 294  
 Db TTGCATTTGAAGGATATGCTCTATTGGATGTGAACCTTGACCCAGGCGGTCTTACCT 964  
 295 Gly 295  
 Db GGG 967

## RESULT 14

CB340603/c 704 bp mRNA linear EST 14-MAR-2003  
 LOCUS CA23BI011VR All Cabernet Sauvignon Leaf - CA23BI Vitis vinifera  
 DEFINITION cDNA clone CA23BI011VR All 3', mRNA sequence.

## ACCESSION

VERSION CB340603.1 GI:28961185

## KEYWORDS

SOURCE Vitis vinifera

## ORGANISM

Vitis vinifera  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 rosids; Vitaceae; Vitis.

## REFERENCE

1. (bases 1 to 704)  
 Goes da Silva, F., Lim, H., Iandolino, A., Baek, J., Leslie, A., Xu, J.,  
 Jones, K., Walker, M.A. and Cook, D.R.

## TITLE

Transcriptional responses of Vitis vinifera to infection by the

bacterial pathogen Xylella fastidiosa

Unpublished (2003)

Contact: Douglas Cook, PhD

CAES Genome Facility

UC Davis, Plant Pathology

One Shields Ave, Davis, CA 95616, USA

Tel: 530 754 6561

Fax: 530 754 6617

Email: drcook@ucdavis.edu

Seq primer: GCCAACGCAATGGCTCTAG.

Location/Qualifiers

1..704

/organism="Vitis vinifera"

/mol\_type="mRNA"

/cultivar="Cabernet Sauvignon"

/db\_xref="taxon:29760"

/clone="CA23BI011VR All"

/sex="Hermaprodite"

/dev\_stage="Mid-season leaf material, collected July 25,

2001."

/lab\_host="DH5alpha"

/clone\_lib="Cabernet Sauvignon Leaf - CA23BI"

/note="Organ: Leaf; Vector: pDNR; Site 1: SfiI; Site 2:

SfiI; CA23BI is a cDNA library of Cabernet Sauvignon

leaves. The leaves were collected on July 25, 2001, in

Napa Valley, California, and represent leaves in

mid-season development. These leaves were verified to be

infected with the bacterial pathogen, Xylella fastidiosa,

based on a diagnostic assay using PCR and Xylella-specific primer pairs. The plants were asymptomatic at the time of collection, but later developed symptoms. cDNAs were made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows: 5'-AACAGTGGTATCAGCAGAGTGGCCATTACGCCGG-3' and 5'-ATTCTAGAGCCGAGCGGCCGACATG-dt(30)NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

## ORIGIN

Alignment Scores: 2,91e-111 Length: 704  
 Pred. No.: 1163.00 Matches: 209  
 Score: 97.4% Conservat: 14  
 Percent Similarity: 97.4% Mismatches: 6  
 Best Local Similarity: 91.3% Indels: 0  
 Query Match: 47.6% Gaps: 0  
 DB: 6

US-10-731-525-8 (1-466) x CB340603 (1-704)

Qy 98 GluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSerGly 117  
 Db GAAATATGATGAAGGCCCTTGCAGAAATTTGGCAATGACATGCCATTGCTTCAGTGGT 644  
 118 AlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArgProPheArgValPhe 137  
 Db GCTGAAGATATTGCTTTGATTGAATATGCACGCTTAACCTGCTCGCTTAGGGTATTC 584  
 138 SerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAspAlaValGluLys 157  
 Db AGCTCGACACTGGAAGGTTGAAACCCGGAACATATCAATTTCTTGACACGGTTGAGAA 524  
 158 HisTyrGlyIleArgIleGluTyrMetPheProAspAlaValGluValGlnAlaLeuVal 177  
 Db CACTATGGCATCCGCATTGAATACATGTTTCCAGATGCTGTTGAAGTTCAAGGCTTG 464  
 178 ArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGluCysCysArgValArg 197  
 Db AGGAGCAAGGAGCTGTTCTCTTTTACGAGGATGGCACCAGGAGTGTCCGCTGTGAGG 404  
 198 LysValArgProLeuArgAlaLeuLysGlyLeuArgAlaTrpIleThrGlyGlnArg 217  
 Db AAGGTGAGACCTTGGAGAGGCACTGAAAGGTTTGGCTGCTGGATCACAGGCGCAAGA 344  
 218 LysAspGlnSerProGlyThrArgSerGluIleProValValGlnValAspProAlaPhe 237  
 Db AAAGATCAGTCTCTGCTGACGGCTGAAGTTTCTGTTGTCAGGTGGATCCAGCCTTC 284  
 238 GluGlyMetAspGlyGlyIleGlySerLeuValIysTrpAsnProValAlaAsnValLys 257  
 Db GAGGGGTGGATGTTGGGTTGGCAGCTGTTGAAATGGAACCCAGTGGCAATGTGCAG 224  
 258 GlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAsnSerLeuHisAla 277  
 Db GGCATGGACATCTGGAATTTCTCTCCGTGCTATGAATGTGCTGTGAATTCATTGCACTG 164  
 278 LysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeuProGlyGlnHis 297  
 Db AAAGGGTATATTTCAATTTGGGTGAGCCATGCACTAGGCTGTATTACTTGGCCAGAT 104  
 298 GluArgGluGlyArgTrpTrpGluAspAlaLysAlaLysGluCysGlyLeuHisLys 317  
 Db GAGAGAGAGGAAGATGTTGGTGGGAAGATGCCAAGGCCAAGGAGTGTGGGCTTCATAAA 44  
 318 GlyAsnValLysGlnLysGluGlu 326  
 Db GGAACCTCAAAACAAAAA 17

## RESULT 15

CK287268

LOCUS

DEFINITION

CK287268 856 bp mRNA linear EST 02-AUG-2004  
 EST749990 Nicotiana benthamiana mixed tissue cDNA library,

normalized, full-length Nicotiana benthamiana cDNA clone NEMB793 5' end, mRNA sequence.

## ACCESSION

CK287268

## VERSION

CK287268.1 GI:39863647

## KEYWORDS

EST.

## SOURCE

Nicotiana benthamiana

## ORGANISM

Nicotiana benthamiana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamiales; Solanales; Solanaceae; Nicotiana.

1 (bases 1 to 856)

## REFERENCE

Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,

## AUTHORS

Skaskawicz, B., Jin, H. and Baker, B.

## TITLE

Generation of EST sequences from Nicotiana benthamiana

## JOURNAL

Unpublished (2003)

## COMMENT

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr. Rockville, MD 20850, USA

Email: potato-array@tigr.org

Clones can be requested from the University of Arizona Genomics

Institute via <http://genome.arizona.edu/orders/>.

Seq primer: ATT TAG GTG ACA CTA TAG.

## FEATURES

source

1..856  
 /organism="Nicotiana benthamiana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4100"  
 /clone="NEMB793"  
 /tissue types="abiotic and biotic stress-treated leaves,  
 callus tissue and root tissue"  
 /lab host="DH10B-TonA"  
 /clone\_lib="Nicotiana benthamiana mixed tissue cDNA  
 library, normalized, full-length"  
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
 supplier: RNA was isolated from Nicotiana benthamiana  
 tissues that include callus, roots from liquid culture  
 grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),  
 cold-stressed leaves (5 C 3 hr, 6hr), and pathogen  
 challenged leaves (Pseudomonas syringae pv tomato 12 hr;  
 Xanthomonas campestris pv campestris 12 hr, 18hr;  
 Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas  
 campestris pv vesicatoria 18hr). RNA was isolated from  
 these tissues and pooled in approximately equal molar  
 amounts."

## ORIGIN

## Alignment Scores:

Pred. No.:	2,11e-110	Length:	856
Score:	1156.00	Matches:	226
Percent Similarity:	88.0%	Conservative:	31
Best Local Similarity:	77.4%	Mismatches:	25
Query Match:	47.3%	Indels:	10
DB:	7	Gaps:	5

US-10-731-525-8 (1-466) x CK287268 (1-856)

Qy	4	AlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSerGluProLysLeu	23
Db	1	GCATTTCATAGCTCTTTGTCG---CCTTCTCTCTTAT-----GAACATCAAGTG	48
Qy	24	ProGlnIleGlySerIleArgIleSerGluArgPro-----IleGlyGlyAlaValAsn	41
Db	49	TCCCAATTGGGGACCTTTTCAGCCATTTGGATAGGCTCAAAATGCTGTCAAGGGCTTTGAAC	108
Qy	42	PheAsnLeuSerGlnArgSerLeuValIysProValAsnAlaGluProProArgLys	61
Db	109	-----ACTCGAGGCGTTCGGCAGTGAACCATTTGAACGCTGACCTTAAGAGGAAT	159
Qy	62	AspSerIleValProLeuAlaAlaThrThrIleValAlaSerAlaSerGlu---ThrLys	80
Db	160	GATTCATAGTTGCTCTGCGCAACCTTCGTGGCTCCTGAAGTAGCAGAGAAACTGAG	219
Qy	81	GluGluAspPheGluGlnIleAlaSerAspLeuAsnAlaSerProLeuGluIleMet	100

Search completed: February 21, 2006, 01:51:34

Job time : 5585 secs

Db	220	GAAGAGGACTTCAGAAATTTGGCTAAGGAGCTTCAACATGCTTCCCTCTTGAGATTATG	279
Qy	101	AspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSerGlyAlaGluAsp	120
Db	280	GACAAGGCCCTTGAGAAATTTGGGGATGACGTTGCTATTCTTTTCAGTGTGCTGAAGAT	339
Qy	121	ValAlaIleGluLeuThrAlaLysLeuThrGlyArgProPheArgValPheSerLeuAsp	140
Db	340	GTAGCTTTGATTGAGTATGCACATTAACTGGTCGACCATTCAGAGTTTTCAGCCITGAT	399
Qy	141	ThrGlyArgLeuAsnProGluThrTyrrGlnLeuPheAspAlaValGluLysHisTyrrGly	160
Db	400	ACTGGGAGGTTGAACCGGAGACCTACCAATTATTTGATAGTGTGGAGAGCATTATGGC	459
Qy	161	IleArgIleGluTyrrMetPheProAspAlaValGluValGlnAlaLeuValArgSerLys	180
Db	460	ATTGCGATTGAGTACATGTTCCCTGATGCTGTTGAAGTTTCAGGCTTAGTAGCCAAA	519
Qy	181	GlyLeuPheSerPheTyrrGluAspGlyHisGlnGluCysCysArgValArgLysValArg	200
Db	520	GGGCTTTTCTCTTCTATGAAGACGGCCACCAAGAGTGTGCCGTGTAAGAGGTTAGG	579
Qy	201	ProLeuArgArgAlaLeuLysGlyLeuArgAlaTrpIleThrGlyGlnArgLysAspGln	220
Db	580	CCCTTAAGAGAGACCCCTCAAGGCTTACGTGCTGGATCACAGGGCAACGTAAGATCAA	639
Qy	221	SerProGlyThrArgSerGluIleProValValGlnValAspProAlaPheGluGlyMet	240
Db	640	TCCCTCGAACTCGATCAGAAATTCGCTGTTTTCAGGTGGACCCCTCTTTTGAAGGTTG	699
Qy	241	AspGlyGlyIleGlySerLeuValIleTrpAsnProValAlaAsnValLysGlyHisAsp	260
Db	700	GATGGCGGCTGCTGGTAGCTTGGTGAAGTGAACCCCTGTGGCTACGTAGACGGAAGAT	759
Qy	261	IleTrpAsnPheLeuArgThrMetAsnValProValAsnSerLeuHisAlaLysGlyTyr	280
Db	760	ATTGGAACTTCTCTCGTGCATGAATGTTCCGGTTAACTCATTCATTACATTCACAAGTTAT	819
Qy	281	ValSerIleGlyCysGluProCysThrArgProVal	292
Db	820	GTCTCCATTGGATCGGAACCTTGCACAAGGCGCGTC	855

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 20, 2006, 22:26:17 ; Search time 322 Seconds  
(without alignments)  
2572.496 Million cell updates/sec

Title: US-10-731-525-8

Perfect score: 2445

Sequence: 1 MALAFSSISAPSTPSSSE.....KYPSEKRDVDSLMFAFNALR 466

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-O=/abs/abssweb.spool/US10731525/runat.17022006.164901.185/app\_query.fasta\_1  
-DB=Issued PatentsE nla -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-HOST=abs06p -USER=US10731525 @CGN 1.1.290 @runat.17022006.164901.185 -NCFU=6  
-ICPU=3 -NO KWAP -NEG SCORES=0 -WAIT -DEBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/1 COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5 COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/H COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/PTUS COMB.seq.\*  
7: /cgn2\_6/ptodata/1/ina/PP COMB.seq.\*  
8: /cgn2\_6/ptodata/1/ina/RE COMB.seq.\*  
9: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2445	100.0	1629	3	US-09-720-318A-7
2	1923	78.7	1795	3	US-09-720-318A-5
3	1645.5	67.3	1827	3	US-09-720-318A-9
4	1295.5	53.0	1210	3	US-09-720-318A-3
5	1262.5	51.6	1215	3	US-09-720-318A-1
6	811.5	33.2	783	3	US-09-328-352-1063
7	745.5	30.5	810	3	US-09-252-991A-13944
8	546.5	22.4	6063	3	US-09-902-540-807
9	540	22.1	723	3	US-09-902-540-7978

10	407.5	16.7	756	3	US-09-540-236-1037	Sequence 1037, Ap
11	405.5	16.6	269223	3	US-09-596-002-41	Sequence 41, Appl
12	307.5	12.6	747	3	US-09-134-001C-1462	Sequence 1462, Ap
13	307.5	12.6	4055	3	US-09-710-279-3357	Sequence 3357, Ap
14	307.5	12.6	4403765	3	US-09-103-840A-2	Sequence 2, Appli
15	307.5	12.6	4411529	3	US-09-103-840A-1	Sequence 1, Appli
16	286	11.7	660	3	US-09-252-991A-13837	Sequence 13837, A
17	266.5	10.9	1240	3	US-09-962-357-7	Sequence 7, Appli
18	250.5	10.2	468	3	US-09-710-279-307	Sequence 307, App
19	209.5	8.6	747	3	US-09-543-681A-2919	Sequence 2919, App
20	202	8.3	783	3	US-09-489-039A-2784	Sequence 2784, Ap
21	168.5	6.9	640681	3	US-09-790-988-1	Sequence 1, Appli
22	157.5	6.4	1664976	3	US-08-916-421B-1	Sequence 1, Appli
23	157.5	6.4	1664976	3	US-09-692-570-1	Sequence 10, Appl
24	156	6.4	2413	3	US-09-613-182-10	Sequence 1, Appli
25	147	6.0	2640	3	US-09-962-357-1	Sequence 1, Appli
26	145	5.9	1547	2	US-08-557-122A-2	Sequence 2, Appli
27	145	5.9	1547	3	US-09-262-666-2	Sequence 2, Appli
28	140	5.7	2170	3	US-09-807-258-9	Sequence 9, Appli
29	139.5	5.7	1101	3	US-08-984-919A-13	Sequence 13, Appl
30	139.5	5.7	1101	3	US-08-984-919A-14	Sequence 14, Appl
31	139.5	5.7	1401	3	US-08-984-919A-32	Sequence 32, Appl
32	139.5	5.7	1401	3	US-08-984-919A-34	Sequence 34, Appl
33	139.5	5.7	1410	3	US-08-984-919A-54	Sequence 54, Appl
34	139.5	5.7	1410	3	US-08-984-919A-56	Sequence 56, Appl
35	139.5	5.7	1466	3	US-08-984-919A-10	Sequence 10, Appl
36	139.5	5.7	1488	3	US-08-984-919A-49	Sequence 49, Appl
37	139.5	5.7	1488	3	US-08-984-919A-50	Sequence 50, Appl
38	139.5	5.7	1518	3	US-10-396-869A-7	Sequence 7, Appli
39	139.5	5.7	1569	3	US-09-368-588-1	Sequence 1, Appli
40	139.5	5.7	1875	3	US-08-984-919A-46	Sequence 46, Appl
41	139.5	5.7	1875	3	US-08-984-919A-48	Sequence 48, Appl
42	139.5	5.7	4040	3	US-09-999-833A-458	Sequence 458, App
43	139.5	5.7	4040	3	US-10-020-445A-458	Sequence 458, App
44	138.5	5.7	1107	3	US-08-781-420-13	Sequence 13, Appl
45	138.5	5.7	1107	3	US-08-781-420-14	Sequence 14, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-720-318A-7  
; Sequence 7, Application US/09720318A  
; Patent No. 6730827  
; GENERAL INFORMATION:  
; APPLICANT: Falco, Saverio Carl  
; APPLICANT: Allen, Stephen M.  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167-C  
; CURRENT APPLICATION NUMBER: US/09/720,318A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 1998-07-14  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 7  
; LENGTH: 1629  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-09-720-318A-7

Alignment Scores:  
Pred. No.: 3 86e-270 Length: 1629  
Score: 2445.00 Matches: 466  
Percent Similarity: 100.0% Conservatives: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 3 Gaps: 0

US-10-731-525-8 (1-466) x US-09-720-318A-7 (1-1629)

QY 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerGlu 20  
|||||



```
Db 32 ATGGCCCTCGCTTCACTTCTCAATTTCCGACCAACTTCCACCTTCCCATCATCGAA 91
Qy 21 ProlylsLeuProGlnIleGlySerIleArgIleSerGluArgProIleGlyGlyAlaVal 40
Db 92 CCCAAATCTCCGAAATGGGTCAATTAAGATTTCGGAGAGGCCCAATTCGAGGCGCGTT 151
Qy 41 AsnPheAsnLeuSerGlnArgSerLeuValLysProValAsnAlaGluProProArg 60
Db 152 AATTTCAATTAATCTCAAGACGGAGCTTGGTAAGCCCGTTAAACGCCGAACTCCACGC 211
Qy 61 LysAspSerIleValProLeuAlaAlaThrThrIleValAlaSerAlaSerGluThrLys 80
Db 212 AAGGATTCATTTGCTCTCGCAGCAACACCATCGTTGCTTCTGCTCTGAGAGAA 271
Qy 81 GluLysPheGluGlnIleAlaSerAspAsnAlaSerProLeuGluIleMet 100
Db 272 GAGGAAGATTTGAACAGATAGCCAGTGATCTGCACAAATGCTTCACCTCTGAAATCATG 331
Qy 101 AspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSerGlyAlaGluAsp 120
Db 332 GATAGAGCCCTCGACAAATTCGGACACGACATGCTATTGGCTTCAGTGTGCTGAAGAT 391
Qy 121 ValAlaLeuIleGluTyrAlaLysLeuThrGlyArgProPheArgValPheSerLeuAsp 140
Db 392 GTTGTCTTGTAGTATCGAAATTCAGCGGTGACCCCTTTAGGGTTTTTCAGTTTTGGAC 451
Qy 141 ThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAspAlaValGluLysHisTyrGly 160
Db 452 ACTGGAGACTGAACCCGAAACTTATCAACTTTTGTATGCGGTGAGAAAGCATTAATGGA 511
Qy 161 IleArgIleGluTyrMetPheProAspAlaValGluValGlnAlaLeuValArgSerLys 180
Db 512 ATTCGCATTGATGATACATGTTCCCTCGATGCTGTTGAGGTTCAGGATTTGGTGAGGTAAG 571
Qy 181 GlyLeuPheSerPheTyrGluAspGlyHisGlnGluCysArgValArgLysValArg 200
Db 572 GGGTTATCTTCTTACGAGGATGGGCACCAAGAGTGTTCACGGTGCAGAAAGGTGAGG 631
Qy 201 ProLeuArgArgAlaLeuLysGlyLeuArgAlaTrpIleThrGlyGlnArgLysAspGln 220
Db 632 CCTTTAAGGAGGGCCCTTAAGGCTCAGAGCATGGATACTTGTCTCAGAGGAAAGACAG 691
Qy 221 SerProGlyThrArgSerGluLeuProValValGlnValLeuAspProAlaPheGluGlyMet 240
Db 692 TCACCTGTGATAGTCTGAAATACCGGTGTTGTTGAGGTGTATCCGGCTTTTTCAGGGAATG 751
Qy 241 AspGlyGlyIleGlySerLeuValLysTrpAsnProValAlaAsnValLysGlyHisAsp 260
Db 752 GATGTTGAATTTGGAGCTTGGTGAAGTGGAAACCTGTTGCAATGTGAAGGGCCATGAC 811
Qy 261 IleTrpAsnPheLeuArgThrMetAsnValProValAsnSerLeuHisAlaLysGlyTyr 280
Db 812 ATATGGAACCTTCTTAGCACCATGAATGTGCTGTGAATTCCTTGCATGCAAAAGGATAT 871
Qy 281 ValSerIleGlyCysGluProCysThrArgProValLeuProGlyGlnHisGluArgGlu 300
Db 872 GTTTCCATTTGGTGTGAGCCCTGCACATAGGCTGTTTACCTGGGCACATGAAGGGAA 931
Qy 301 GlyArgTyrTrpTyrGluAspAlaLysAlaLysGluCysGlyLeuHisLysGlyAsnVal 320
Db 932 GGGAGTGTGTGGTGGGAGGATGCCAAGCTAAGGNAATGTGCTTCAAAAGGAAATGTA 991
Qy 321 LysGlnGlnLysGluGluAspValAsnGlyAsnGlyLeuSerGlnSerHisAlaAsnGly 340
Db 992 AAGCAGCAGAAAGAGGAGGATGTTAATGCAAAATGGGCTATCCCAATCCATGCAAAATGGT 1051
Qy 341 AspAlaThrThrValProAspIlePheAsnSerProAsnValValAsnLeuSerArgThr 360
Db 1052 GATGCTACCTGTGCTCCCTGACATTTTCAACAGCCCGCAATAGTATTAATCTGAGCAGGACT 1111
Qy 361 GlyIleGluAsnLeuAlaLysLeuGluAspArgLysGluProTrpLeuValValLeuTyr 380
Db 1112 GGAATTTGAGAAATTTGGCAAAATTTGAGAGCCGGAAGGAAACCATGGCTTGTGCTTTAT 1171
```

```
Qy 381 AlaProTyrCysProTyrCysGlnAlaMetGluGluSerTyrValAspLeuAlaAspLys 400
Db 1172 GCACCAATGGTGCCCTACTGCCAGGCTATGGAGGAATCTTATGTTGACTTAGCAGACAAG 1231
Qy 401 LeuAlaGlySerThrGlyMetLysValGlyLysPheArgAlaAspGlyGluGlnLysGlu 420
Db 1232 TTAGCAGGGTCAACAGCGATGAAGTTTGGAAATTTAGAGCAGATGGAGAACAGAAAGAA 1291
Qy 421 PheAlaLysSerGlnLeuGlnLeuGlySerPheProThrIleLeuPhePheProLysHis 440
Db 1292 TTTGCAAAAGAGTGAATGCAATTTGGGAAGCTTCCCTACGATATATTTTCCCAAGCAT 1351
Qy 441 SerSerArgProThrIleLysTyrProSerGluLysArgAspValAspSerLeuMetAla 460
Db 1352 TCGTCTCGCCACACATAAAGTATCCCTCAGAAAGAGAGATGTTGATTCCTTGTATGCA 1411
Qy 461 PheValAsnAlaLeuArg 466
Db 1412 TTTGTAATGCCCTTAAGA 1429

RESULT 2
US-09-720-318A-5
; Sequence 5, Application US/09720318A
; Patent No. 6730827
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/09/720,318A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1795
; TYPE: DNA
; ORGANISM: Glycine max
US-09-720-318A-5

Alignment Scores:
Pred. No.: 3,64e-210 Length: 1795
Score: 1923.00 Matches: 378
Percent Similarity: 87.2% Conservatives: 37
Best Local Similarity: 79.4% Mismatches: 45
Query Match: 78.7% Indels: 16
DB: 3 Gaps: 7

US-10-731-525-8 (1-466) x US-09-720-318A-5 (1-1795)
Qy 1 MetAlaLeuAlaPheThrSerSerIleSer-----AlaProThrSerThr 15
Db 56 CTTGCCGCTTCCACTACTTCTTCTCTCAGTCAGCAGCAGCAGCGCTCGAGCTCT 115
Qy 16 Phe-----ProSerSerGluProLysLeuProGlnIleGlySerIleArg 30
Db 116 TTCTTCTCGCGCTTGGATCTTCAATCGGAGCTAAAGCTCCGCAATTTGGTTCCTTCGG 175
Qy 31 IleSerGluArgProIleGlyGlyAlaValAsnPheAsnLeuSerGlnArgSerLeu 50
Db 176 TTTCGGAGAGCCCTCAAGTTTCTGCTGGTGTGTTTAAATTAATCAAGACGCTCCTCG 235
Qy 51 ValLysProValAsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThr 70
Db 236 GTGAGGCCACTCAATCGCAACCCGAAACCGCAATGATTTCTTGTGTTCTCTTTCGACGA--- 292
Qy 71 ThrIleValAlaSerAlaSerGluThrLysGluGluAspPheGluGlnIleAlaSerAsp 90
Db 293 ACTATCGTTGCTCTCGAGGTTGAGAGAGAGAAAGATTTTGTAGCAATTTAGCGAAGAC 352
Qy 91 LeuAspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAsp 110
```





```
Qy 195 ArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaTrpIleThr 214
Db 680 AGGTGAGGAGGTTGGCCCTTGAGAGGGCCCTCAAGGGCTCAAGGGCTGGATCACC 739
Qy 215 GlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValValGlnValAsp 234
Db 740 GGGCAGAGGAGGATCAGTCCCTGGCCACAGAGCCAGCATCCCTGTGTTCAGTTGAT 799
Qy 235 ProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsnProValAla 254
Db 800 CCGTCTTTTGAAGGGCTGGATGGTGGAGCCGGTGGTTCATCAAGTGAACCCCTGTGGCT 859
Qy 255 AsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAsnSer 274
Db 860 AATGTGGATGGCAAGGATATCTGGACCTTCTCAGGACCATGGATGTCCTGTGGACACC 919
Qy 275 LeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeuPro 294
Db 920 CTGCATGCTCAAGGCTAGCTCCATTTGGTGGAGCCGTGCACAGGCCCGTGTGGCG 979
Qy 295 GlyGlnHisGluArgGluGlyArgTrpTrpTrpGluAspAlaLysAlaLysGlyCysGly 314
Db 980 GGGCAGCAGCAGAGGAGGAGGTGTGTGGAGGACGCCACGGCCAAAGGAGTGGCGC 1039
Qy 315 LeuHisLysGlyAsnValLys-----GlnGlnLysGluGluAspValAsnGlyAsn 331
Db 1040 CTGCACAACTGATACATCGACAGGAAGGTTCAGGCCACCAAGTCCGCGTCAACGGCAAC 1099
Qy 332 GlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSer 351
Db 1100 GGCTCGCCGAGCCAGCGCC-----CCAGACATCTTCCAGAGC 1138
Qy 352 ProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArg 371
Db 1139 CAGGCCCATCTGTCACCTCACCCTCCCGGGATCGAGAACCTCTCCGCGCTCGAGAACCGC 1198
Qy 372 LysGluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMetGlu 391
Db 1199 GCGAGCCGCGTGCCTACCGTCTCTACGCTCCCTGGTGGCCATCTCCAGGCAATGGAG 1258
Qy 392 GluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLys 411
Db 1259 GCGTCTCTACGTGAGCTGGCCGAGAGCTGAGCGGCTCA--GGCATCAAGGTGGCCAAAG 1315
Qy 412 PheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPhe 431
Db 1316 TTCCGCGCGCAGCGCAGAGCCATTCGCGCAGCGGAGCTGCAACTACAGAGCTTC 1375
Qy 432 ProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSerGlu 451
Db 1376 CCGACGATCTCTGTTCCCGCGCGCCACCGTGAAGGCC--ATCAAGTACCCGCTCCGAG 1432
Qy 452 LysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
Db 1433 AAGAGGACGTCAGCTCCCTCTCCCTCTGCTGAGAACAGCCCTCAGA 1477
```

## RESULT 4

```
US-09-720-318A-3
; Sequence 3, Application US/09720318A
; Patent No. 6730827
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/09/720,318A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 1210
```

```
; TYPE: DNA
; ORGANISM: Impatiens balsamia
US-09-720-318A-3

Alignment Scores:
Pred. No.: 1,95e-138 Length: 1210
Score: 1295.50 Matches: 239
Percent Similarity: 88.1% Conservative: 28
Best Local Similarity: 78.9% Mismatches: 29
Query Match: 53.0% Indels: 7
DB: Gaps: 3
```

US-10-731-525-8 (1-466) x US-09-720-318A-3 (1-1210)

```
Qy 165 TyrMetPheProAspAlaValGluValGlnAlaLeuValArgSerLysGlyLeuPheSer 184
Db 9 TACATGTTCCCTGATGCAATTTAGTACAGGATTAGTAAAGAACCAAGACATGTTCTCT 68
Qy 185 PheTyrGluAspGlyHisGlnGluCysArgValArgLysValArgProLeuArgArg 204
Db 69 TTCTACGAGACCGACATCAAGAGTGTCTGCGCGTCAGAAAAGTGAGGCCACTGAGCGT 128
Qy 205 AlaLeuLysGlyLeuArgAlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThr 224
Db 129 GCTCTCAAGGGTCTCCGCGCTTCGATCACGCGGCAAAAGAACCAAGTCCCGGGAACG 188
Qy 225 ArgSerGluIleProValValGlnValAspProAlaPheGluGlyMetAspGlyGlyLe 244
Db 189 AGATCGAGAGATCCAGTCTGTCAGGTGATGATCCCTCTTTGAGGGATGCTTGTGGAGAG 248
Qy 245 GlySerLeuValLysTrpAsnProValAlaAsnValLysGlyHisAspIleTrpAsnPhe 264
Db 249 GGTAGCTGTGTAAGTGGATCCGCTGGCTAATGTAGATGCTGCTGATGATGATGAATTC 308
Qy 265 LeuArgThrMetAsnValProValAsnSerLeuHisAlaLysGlyTyrValSerIleGly 284
Db 309 CTCGAGCTATGAATGCTGCTTAATGCACTTCATAGCCAGGGTTATGTCTCGAATGGG 368
Qy 285 CysGluProCysThrArgProValLeuProGlyGlnHisGluArgGluGlyArgTrp 304
Db 369 TGCAGAACCGTGCACCGCGGTGTACCTGGGCAACATGAGAGAGAGGAGGAGTGTGG 428
Qy 305 TrpGluAspAlaLysAlaLysGluCysGlyLeuHisLysGlyAsnValLysGlnGlnLys 324
Db 429 TGGAGAGATCTCGCGCTAAGAGTGTGGCTACATAAAGGAAATATAAAG----- 479
Qy 325 GluGluAspValAsnGlyAsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThr 344
Db 480 -----GATGCCAATGGGAATGGGTGCTCAAGCTGAGGGAGGAGGAAGAACTGTTACG 533
Qy 345 ValProAspIlePheAsnSerProAsnValValAsnLeuSerArgThrGlyIleGluAsn 364
Db 534 GATGCTGATATTTTGAATCCAGAAATGGTGACACTGAGTAGAGAGGAGGATTTGAGAAT 593
Qy 365 LeuAlaLysLeuGluAspArgLysGluProTrpLeuValValLeuTyrAlaProTrpCys 384
Db 594 CTGTGGAATCTCAGGAGAGGAAAGACCATGATCGTGTGCTGTATGACACTTGTGTGC 653
Qy 385 ProTyrCysGlnAlaMetGluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySer 404
Db 654 CAGTTCTGCCAGGATATGGAATAATCATCTACTTGAATTTGGCTGAAAAAGCTGGCGGTGAGC 713
Qy 405 ---ThrGlyMetLysValGlyLysPheArgAlaAspGlyGluGlnLysGluPheAlaLys 423
Db 714 GTGCGTGTGTGAAGGTAGGAAATCCGGGCAGATGTCAGAAAAGAGATTGTCTCAC 773
Qy 424 SerGluLeuGlnLeuGlySerPheProThrIleLeuPhePheProLysHisSerSerArg 443
Db 774 CAAGAATTCAGCTGGCGGAGCTTTCCAAACAATCTCTTCTCCCAAAACACTCATCTAAA 833
Qy 444 ProThrIleLysTyrProSerGluLysArgAspValAspSerLeuMetAlaPheValAsn 463
Db 834 ---GCCATCAAGTACCCATCTGAGAAAAGGAGCGTGAGTCATTGTTGGCTTTTGTGAAC 890
```

```
Qy 464 AlaleuArg 466
Db 891 GCACTCAGA 899

RESULT 5
US-09-720-318A-1
; Sequence 1, Application US/09720318A
; Patent No. 6730827
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/09/720,318A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1215
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (273)
; OTHER INFORMATION: n = a, c, g or t
US-09-720-318A-1

Alignment Scores:
Pred. No.: 1,2e-134 Length: 1215
Score: 1262.50 Matches: 235
Percent Similarity: 86.2% Conservative: 27
Best Local Similarity: 77.3% Mismatches: 27
Query Match: 51.6% Indels: 15
DB: 3 Gaps: 4

US-10-731-525-8 (1-466) x US-09-720-318A-1 (1-1215)

Qy 167 PheProAspAlaValGluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyr 186
Db 3 TTCCCGGACCGGACGAGGTCGAGGCTGGTGGCCACCAAGGCGCTTCTCTCTCTAC 62
Qy 187 GluAspGlyHisGlnGluCysCysArgValArgLysValArgProLeuArgAlaLeu 206
Db 63 GAGAGCGCCACCAAGAGTCTGCGGTGGCCAGGTGGCGCCCTGCGCAGGGGCTC 122
Qy 207 LysGlyLeuArgAlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSer 226
Db 123 AAGGGGCTTAGGGCATGGATCACCAGGCGCAGAGGAAGACCAAGTCCCGGCGCAGGGCC 182
Qy 227 GluIleProValGlnValAlaAspProAlaPheGluGlyMetAspGlyIleGlySer 246
Db 183 AGATCCCATTTGTCAGGTGTGATCTCTTCCGAAGGCTGGATGGCGGGCGGTAGC 242
Qy 247 LeuValLysTrpAsnProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArg 266
Db 243 TTGGTCAGTGGAAACCCGCGGCGCCAGTCNACGGCAGAGCATCTGGACTTCTTGGCG 302
Qy 267 ThrMetAsnValProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGlu 286
Db 303 ACCATGGACGTACTGTCAACACCTGCTCAGGGCTACGTGTCCATCGGGTGGAG 362
Qy 287 ProCysThrArgProValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpTrpGlu 306
Db 363 CCGTGCAACAGGCGCGTCTGCGGGGCGCAGCAGCGTGAAGCGCGGTGGTGGAG 422
Qy 307 AspAlaLysAlaLysGluCysGlyLeuHisLysGlyAsnValLysGlnGlnLysGlu--- 325
Db 423 GACGCCNAGCCAGGAGTGGCGCTCCACAGGGCCAACTTACNAGGAGCGGCCAGGCG 482
Qy 326 -----GluAspValAsnGlyAsnGlyLeuSerGlnSerHisAlaAsnGlyAspAla 342

483 GCGGCCCCCAGGTCCGCCAACGCGAACGCG-:::TCG 515
343 ThrThrValProAspIlePheAsnSerProAsnValValAsnLeuSerArgThrGlyLe 362
516 GCGGGGCGCCCGGACATCTTCGAGAGCCCGCGCTGGTGTCTCCACCCGACCGGATC 575
363 GluAsnLeuAlaLysLeuGluAspArgLysGluProTrpLeuValValLeuTyrAlaPro 382
576 GAGAACCTGCTGCGCTGGAGAACCGCGCGAGCGCTGGTGTCTGTGCTGTACGCGGCC 635
383 TrpCysProTyrCysGlnAlaMetGluGluSerTyrValAspLeuAlaAspLysLeuAla 402
636 TGGTGGCGGCTTTCGCCAGGCGCATGGAGCGCTCTAGCTGGAGCTGGCGGAGAGCTGGCG 695
403 GlySerThrGlyMetLysValGlyLysPheArgAlaAspGlyGluGlnLysGluPheAla 422
696 GGGTCC---GGGTGAAGGTGGCCAAAGTTCGCGCGGACGCGGAGCAGAAAGCGCTTCGCG 752
423 LysSerGluLeuGlnLeuGlySerPheProThrIleLeuPhePheProLysHisSerSer 442
753 CAGGCGGAGCTGCAGCTGCAGAGCTTCCACCGTGTCTCTGTTCCGCGCGCGCACCGCC 812
443 ArgProThrIleLysTyrProSerGluLysArgAspValAspSerLeuMetAlaPheVal 462
813 AGGCC---ATCAAGTACCGTCGAGAGAGGAGCGTCCGACTCGCTCTCTGCTTCGCTTCGTC 869
463 AsnAlaLeuArg 466
870 AACAGCTTCGG 881

RESULT 6
US-09-328-352-1063
; Sequence 1063, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1063
; LENGTH: 783
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-1063

Alignment Scores:
Pred. No.: 3,34e-83 Length: 783
Score: 811.50 Matches: 150
Percent Similarity: 72.9% Conservative: 41
Best Local Similarity: 57.3% Mismatches: 62
Query Match: 33.2% Indels: 9
DB: 3 Gaps: 3

US-10-731-525-8 (1-466) x US-09-328-352-1063 (1-783)

Qy 61 LysAspSerIleValProLeuAlaAlaThrIleValAlaSerAlaSerGluThrLys 80
Db 22 CGCCCAAGCGCTTTGGCAGGACACCATGACCGTTATTCGACTATT----- 69
Qy 81 GluLysAspPheGluGlnIleAlaSerAspLeuAspAsnAlaSerProLeuGluIleMet 100
Db 70 ---GATATCGTTGATCTCTCGCGCTGAATATGACAGTAAAGTCCGCGAGATTCTA 126
101 AspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSerGlyAlaGluAsp 120
127 GAACCTGCAATTAAGCCACCAAGGT---GAGATCGCAATTTCGTTCTCTGGAGCAGAAGAT 183
121 ValAlaLeuIleGluTyrAlaLysLeuThrGlyArgProPheArgValPheSerLeuAsp 140
```

Db		184	GTAGTCTTAATTGATATACGGCTTCACGTTTAGTGAACCTTCCGTGTTTTAGTCTTGAT	243
Qy		141	ThrGlyArgLeuAsnProGluThrTyrglnLeuPheAspAlaValGluLysHisIstyrGly	160
Db		244	ACAGCGCGTCTACACCCTGAAACTTATCAATTTATTGAAACTGTTCGTAAGCACTACAAT	303
Qy		161	IleArgIleGluTyrMetPheProAspAlaValGluValGlnAlaLeuValArgSerLys	180
Db		304	ATAAATATCGAGATTGTTCCTCGATCGGAAGCTGTGCAAAGCATGGTGAATGAAAAA	363
Qy		181	GlyLeuPheSerPheTyrGluAspGlyHisGlnGluCysArgValArgLysValArg	200
Db		364	GGCTATTATTAGTTTCTTTAAGATGGACATCAAGATGCTGTGCATACGAAAAGTTCAA	423
Qy		201	ProLeuArgArgAlaLeuLysGlyLeuArgAlaTrpIleThrGlyGlnArgLysAspGln	220
Db		424	CCTTTACGTAAAGAACTCGCGACCTTAGATGGTTGGATTACAGGTACGCGTAAAGATCAA	483
Qy		221	SerProGlyThrArgSerGluIleProValValGlnValAspProAlaPheGluGlyMet	240
Db		484	AGCCCTGGTACACGTACGGAATTCAGTGTGTACAGCTGATGACGGCTTCTCTGGTCCA	543
Qy		241	AspGlyGlyIleGlySerLeuValLysIstirPasnProValAlaAsnValLysGlyHisAsp	260
Db		544	GGT-----AAGCNACTCATTAATAATATCGCTTGGCACTGGTCTAGCGCCGAT	594
Qy		261	IleTrpAsnPheLeuArgThrMetAsnValProValAsnSerLeuHisAlaLysGlyTyr	280
Db		595	GTGTGGAGTTATATCCGATGATGAAATTCATATAACCCCTCTACACGAGCGTGGTTTT	654
Qy		281	ValSerIleGlyCysGluProCysThrArgProValLeuProGlyGlnHisGluArgGlu	300
Db		655	GTTCGATTGGTGTGAGCCATGTACAGTCCAGTACTTCCAAATCAACATGAACGTGAA	714
Qy		301	GlyArgTrpTrpGluAspAlaLysAlaLysGluCysGlyLeuHisLysGlyAsnVal	320
Db		715	GGCGGTGGTGGTGGAGAACAACAAAAAGAAATGTGCTTACATGCAGGTAACTTG	774
Qy		321	LysGln	322
Db		775	AAAAAA	780

## RESULT 7

```

US-09-252-991A-13944
; Sequence 13944, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 3312
; SEQ ID NO 13944
; LENGTH: 810
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13944

```

Alignment Scores:		
Pred. No.:	1,32e-75	Length:
Score:	745.50	Matches:
Percent Similarity:	73.9%	Conservative:
Best Local Similarity:	54.6%	Mismatches:
Query Match:	30.5%	Indels:
DB:	3	Gaps:
		1
		810

US-10-731-525-8 (1-466) X US-09-252-991A-13944 (1-810)

Qy 83 AspPheGluGlnIleAlaSerAspLeuAspLeuSerProLeuGluIleMetAspArg 102  
 Db 91 GACCTTCCCGCGCTTCAGCGCTTCCGACAAATCCCGCCAGGACATCTCGAAAGCC 150  
 Qy 103 AlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSerGlyAlaGluAspValAla 122  
 Db 151 GCCTTGCGGACACTTCGGGACGAGCTGTGGATCTCCTTCAGCGGCGCCGAGGACGTGCTC 210  
 Qy 123 LeuIleGluTyrAlaLysLeuThrGlyArgProPheArgValPheSerLeuAspThrGly 142  
 Db 211 CTAGTAGACATCGGCTTGGAGACTCAACCCGCAACGTGAAGGTGTTTCAGGCTCGACACCGGT 270  
 Qy 143 ArgLeuAsnProGluThrTyrGlnLeuPheAspAlaValGluLysHisTyrGlyIleArg 162  
 Db 271 CGCCTGCACCGCGAAACCTATCTGTTTCATCAGCAGGTCGCGCAACATCAGCGCATCGCC 330  
 Qy 163 IleGluTyrMetPheProAspAlaValGluValGlnAlaLeuValArgSerLysGlyLeu 182  
 Db 331 ATCAGCTGCTTTCCCGGACCCCGCGCTGCTCGAACCCCTGGTGAAGGAAAAGGCGCTG 390  
 Qy 183 PheSerPheTyrGluAspGlyHisGlnGluCysArgValArgLysValArgProLeu 202  
 Db 391 TTCAGCTTCTACCGGACGGCCACCGGTGAGTGCTCGCGCATCCGCAAGATCGAAACCGCTC 450  
 Qy 203 ArgArgAlaLeuLysGlyLeuArgAlaTrrPileThrGlyGlnArgLysAspGlnSerPro 222  
 Db 451 AAGCGCAAGCTTCGCGCGGTGCGCGCTTGGGCCACCGGCCCAACGCGCGACCGAGCGCC 510  
 Qy 223 GlyThrArgSerGluIleProValValGlnValAspProAlaPheGluGlyMetAspGly 242  
 Db 511 GGCACGGCAGCCAGGTGGCGTGTCTGGAATCGAGCGGTGCTTTTCCACGCGCGAA--- 567  
 Qy 243 GlyIleGlySerLeuValLysTrrAsnProValAlaAsnValLysGlyHisAspIleTrr 262  
 Db 568 -----AAGCGCGTGTACAAATTCAAACCCACTGTCTCATGACCAGCGAGGAGGTCTGG 621  
 Qy 263 AsnPheLeuArgThrMetAsnValProValAsnSerLeuHisAlaLysGlyTyrValSer 282  
 Db 622 GGCTATATCCGCATGCTCGAACTGCGCTTACAAACAGCCTGCACGAAACGCGGCTATATCAG 681  
 Qy 283 IleGlyCysGluProCysThrArgProValLeuProGlyGlnHisGluArgGluGlyArg 302  
 Db 682 ATCGGCTGCGAAACCTTGACCCGTCGGTCCGTGTCGCCAACACGACGAGCGCGAGGCGCC 741  
 Qy 303 TrrTrrTrrGluAspAlaLysAlaLysGluCysGlyLeuHisLysGlyAsnVal 320  
 Db 742 TGGTGGTGGGGAAGCAACCCCAAGGAGTGGCGGTGTCACCGCGCAACCTG 795

RESULT 8  
 US-09-902-540-807/c  
 ; Sequence 807, Application US/09902540  
 ; Patent No. 6833447  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Wiegand, Roger C.  
 ; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
 ; FILE REFERENCE: 38-10(15849)B  
 ; CURRENT APPLICATION NUMBER: US/09/902,540  
 ; CURRENT FILING DATE: 2001-07-10  
 ; PRIOR APPLICATION NUMBER: 60/217,883  
 ; PRIOR FILING DATE: 2000-07-10  
 ; NUMBER OF SEQ ID NOS: 16825  
 ; SEQ ID NO 807  
 ; LENGTH: 6063  
 ; TYPE: DNA  
 ; ORGANISM: Myxococcus xanthus  
 US-09-902-540-807

Alignment Scores: 2.49e-51 Length: 6063  
 Pred. No.:

Score: 546.50 Matches: 118  
 Percent Similarity: 59.3% Conservative: 45  
 Best Local Similarity: 42.9% Mismatches: 81  
 Query Match: 22.4% Indels: 31  
 DB: 3 Gaps: 9

US-10-731-525-8 (1-466) x US-09-902-540-807 (1-6063)

Qy 46 GlnArgSerLeuValProValAsnAlaGluProProArgLysAspSerIle 64  
 Db AAGAGCGGGCTTTCTTCTCGAGTACCAGTCCGCCCGCAGCCCCAG-----3824

Qy 65 ValProLeuAlaThrThrIleValAlaSerAlaSerGluThrLysGluGluAspPhe 84  
 Db -----GCTCGGCCCTCTCGCGGAGCGCTGAAGTGAAGGAGCGCCCC---3779

Qy 85 GluGlnIleAlaSerAspLeuAspAsnAlaSerProLeuGluIleMetAspArgAlaLeu 104  
 Db -----GCGAGGACCTC-----CTCGCTGGACCGAGCGC-----3749

Qy 105 AsplysPheGlyAsnAspIleAlaIleAlaPheSer---GlyAlaGluAspValAlaLeu 123  
 Db -----CGCTTCGTCGCGCGCGCCATCGCTCCAGCTTCGCGTGGAGGACATGGTCTC 3692

Qy 124 IleGluThrAlaLysLeuThrGlyArgProPheArgValPheSerLeuAspThrGlyArg 143  
 Db ATCGACCTGGCGCGCGAGCATGCGCCCGAGCGCTGCGCTTCACGCTCGACACCGGACGC 3632

Qy 144 LeuAsnProGluThrThrGlnLeuPheAspAlaValGluLysHisThrGlyIleArgIle 163  
 Db CTGCCCCGAGAGCATGAGAACTCATGGAGGTGGTGAAGCGCTACGGCGTACCCGTG 3572

Qy 164 GluThrMetPheProAspAlaValGluValGlnAlaLeuValArgSerLysGlyLeuPhe 183  
 Db GAGAGCTACTTCCCGAGCGCGCGCGTGGAGCGTGGAGTCCACGAGCGGTACTTTC 3512

Qy 184 SerPheThrGlu-----AspGlyHisGlnGluCysCysArgValArgLysValArgPro 201  
 Db TCCTTCGCGCAGAGCCTTAGAGCAGCAGCAAGCGCTGCTGCGCCATCCGCAAGGTGGAGCCC 3452

Qy 202 LeuArgAlaLeuLysGlyLeuArgAlaThrIleThrGlyGlnArgLysAspGlnSer 221  
 Db CTGTGCGCGCGCTCGCGGCTCAACAGCGTGGGTGACGGGATTCGCGCGTGGAGATCC 3392

Qy 222 ProGlyThrArgSerGluIleProValValGlnValAspProAlaPheGluGlyMetAsp 241  
 Db -----GGGCTGCTCAAGCTCAACCGCTGGCCACGTAGAGTGGACAGCGGCAT-----3347

Qy 242 GlyGlyIleGlySerLeuValLysTrpAsnProValAlaAsnValLysGlyHisAspIle 261  
 Db -----GGGCTGCTCAAGCTCAACCGCTGGCCACGTAGAGTGGACAGCGGCATC 3299

Qy 262 TrpAsnPheLeuArgThrMetAsnValProValAsnSerLeuHisAlaLysGlyThrVal 281  
 Db TGGGCTTACCTCGCGGCGAAGCGGTGCTTACACGCGCTCCATGACCGTGGCTACCCG 3239

Qy 282 SerIleGlyCysGluProCysThrArgProValLeuProGlyGlnHisGluArgGluGly 301  
 Db TCATCGGCTGTGGCCCTGACCGCGCGCGGTGAACCCCTACGAGNACGAGCGCGCGGC 3179

Qy 302 ArgTrpTrpTrpGluAspAlaLysAlaLysGluCysGlyLeuHis 316  
 Db CGCTGGTGGTGGAGTCCGCGGAGAACCGCGAGTGGCGGCTCCAC 3134

## RESULT 9

US-09-902-540-7978  
 ; Sequence 7978, Application US/09902540  
 ; Patent No. 6833447  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Goldman, Barty S.  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
 ; FILE REFERENCE: 38-10(15849)B  
 ; CURRENT APPLICATION NUMBER: US/09/902,540  
 ; CURRENT FILING DATE: 2001-07-10  
 ; PRIOR APPLICATION NUMBER: 60/217,883  
 ; PRIOR FILING DATE: 2000-07-10  
 ; NUMBER OF SEQ ID NOS: 16825  
 ; SEQ ID NO 7978  
 ; LENGTH: 723  
 ; TYPE: DNA  
 ; ORGANISM: Myxococcus xanthus  
 ; US-09-902-540-7978

Alignment Scores: 4.2e-52 Length: 723  
 Pred. No.: 540,00 Matches: 114  
 Score: 59.9% Conservative: 43  
 Best Local Similarity: 43.5% Mismatches: 75  
 Query Match: 22.1% Indels: 30  
 DB: 3 Gaps: 8

US-10-731-525-8 (1-466) x US-09-902-540-7978 (1-723)

Qy 58 ProProArgLysAspSerIleValProLeuAlaAlaThrThrIleValAlaSerAlaSer 77  
 Db 7 CCCCCGAGCCCCAG-----GCCTCGCCCTGCTCGCGGAGCGCGCT 48

Qy 78 GluThrLysGluGluAspPheGluGlnIleAlaSerAspLeuAspAsnAlaSerProLeu 97  
 Db 49 GAAGTGAAGAGACCCCC-----GCCAGGACCTC-----CTC 81

Qy 98 GluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSer--- 116  
 Db 82 GCCTGGACCGAGCGC-----CGCTTCGTCGCGCGCGCGCATCGCTCCAGCTTC 132

Qy 117 GlyAlaGluAspValAlaLeuIleGluThrAlaLysLeuThrGlyArgProPheArgVal 136  
 Db 133 GCGCTGGAGGACATGCTCTCATCGACCTGGCGCGCCAGCATCGCGCCAGCTCGCGCTC 192

Qy 137 PheSerLeuAspThrGlyArgLeuAsnProGluThrThrGlnLeuPheAspAlaValGlu 156  
 Db 193 TTCACCTGACACCGAGCGCTGCCCCGAGGACGTACGACTCATGAGGTGGTGGCT 252

Qy 157 LysHisThrGlyIleArgIleGluThrMetPheProAspAlaValGluValGlnAlaLeu 176  
 Db 253 AAGCGCTACGGCTCACCGTGGAGAGCTACTTCCCGAGCGCGCGCGCGGTG 312

Qy 177 ValArgSerLysGlyLeuPheSerPheThrGlu-----AspGlyHisGlnGluCysCys 194  
 Db 313 GAGTCCACGAAACGGCTACTTCTCTCCCGAGAGCTAGAGGACGCAAGGCGTGTGC 372

Qy 195 ArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaThrIleThr 214  
 Db 373 GCCATCCGAAAGTGGAGCCCCCTGTCGCGCGCTCGCGGTCAACAGGCGTGGTACG 432

Qy 215 GlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValValGlnValAsp 234  
 Db 433 GGATTGGCGCGTGGAGCGCTCCGTC---ACCCGACCGAGCTGGCGAGCTTAGAGGTGAC 489

Qy 235 ProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsnProValAla 254  
 Db 490 AGCGCGCAT-----GGGCTGCTCAAGCTCAACCGCGTGGCC 525

Qy 255 AsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAsnSer 274  
 Db 526 ACCTGAGGAGCGCGGAGCATCTGGGCTAGTCCGGCGGAGAGCGGTGCCCTCAACGCG 585

Qy 275 LeuHisAlaLysGlyThrValSerIleGlyCysGluProCysThrArgProValLeuPro 294  
 Db 586 CTCCATGACCGTGGCTACCGTCCATCGCTGTCGCGCTGTCACGCGCGCGGTGAAACCC 645

Qy 295 GlyGlnHisGluArgGluGlyArgThrTrpTrpGluAspAlaLysAlaLysGly 314  
 Db 314





QY 263 AsnPheLeuArgThrMetAsnValProValAsnSerLeuHisAlaLysGlyTyrValSer 282  
Db 233420 GCATATATTTGACCAAAATATACCGTTTAATGAGCTGATCACCAAGGCTATCCTTCC 233479  
QY 283 IleGlyCysGluProCysThrArgProValLeuProGlyGlnHisGluArgGluGlyArg 302  
Db 233480 ATTGGCTGTGAGCCTTGTACCATGCCTGTCAAGCAAGCGGAGGATATCCGTGCAGGACGC 233539  
QY 303 TrpTrpTrpGluAspAlaLysAlaLysGluCysGlyLeuHisLys 317  
Db 233540 TGGTGGTGGGAGCATAAAGATATAAAGAGTGGCGTTTACACAAA 233584  
RESULT 12  
US-09-134-001C-1462  
; Sequence 1462, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; PRIOR FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 1462  
; LENGTH: 747  
; TYPE: DNA  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-1462  
Alignment Scores:  
Pred. No.: 2,13e-25 Length: 747  
Score: 307.50 Matches: 75  
Percent Similarity: 51.4% Conservative: 39  
Best Local Similarity: 33.8% Mismatches: 91  
Query Match: 12.6% Indels: 17  
DB: 3 Gaps: 8  
US-10-731-525-8 (1-466) x US-09-134-001C-1462 (1-747)  
QY 98 GluileMetAspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSer--- 116  
Db 106 GAGATTTTAAATGGGCATATCAAACTATGAAATGATATGTTTATCTTCAGTTT 165  
QY 117 GlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArgProPheArgVal 136  
Db 166 GGTGCTGAGAGTATGCTTTTGTGATTGACTTCTCAAAATTAACCCGACGACAAAT 225  
QY 137 PheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAspAlaVal--- 155  
Db 226 GTATTTTATAGACACTGATTACATTTTCAGAAACATATGACTTAATAGTAGGGTAA 285  
QY 156 GluLysHisTyrGlyIleArgIleGluTyrMetPheProasp---AlaValGluValGln 174  
Db 286 GATAAATATCCGAATATACGATTAATAAATGAAAAACCCGAACTTACGTAGAGGAA 345  
QY 175 AlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGluCys 194  
Db 346 GGTGAGAAATATAATCCTGCTTTA-----TGAAGAATGATCCTAACCAATGTTGC 396  
QY 195 ArgValArgLysValArgProLeuArgAlaLeuLysGlyLeuArgAlaTrpIleThr 214  
Db 397 TACATACGCAAGATTAACACCACTAGAGACGTATATCTGGTCTGTAGCTGGATATCA 456  
QY 215 GlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValGlnValAsp 234  
Db 457 GGTCTTAGAGCAGACAAATCACCA---ACACGAGCAGATACAAATTTCAATACAAAGAT 513  
QY 235 ProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsnProValAla 254

Db 514 GAAAGATTTAAGTCAATATAA-----GTGTGTCCTTAATCTATTGGACAGAA----- 561  
QY 255 AsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAsnSer 274  
Db 562 -----GAAGAAGTATGCTTATATACGTGATAGGACTTACCTTTAATGAA 609  
QY 275 LeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeuPro 294  
Db 610 TTACATGATCAAAATTTATCAAGTATTTGTTGCAATTCATGTACATCACCCTGTTTGTAT 669  
QY 295 GlyGlnHisGluArgGluGlyArgTrpTrpGluAspAlaLysAlaLysGlyCysGly 314  
Db 670 TCTAATGATTCAAGTCTGCTGCTGCT-----TGTGTCATTTCTAGTAAGACTGAATGCGGA 723  
QY 315 LeuHis 316  
Db 724 TTACAT 729  
RESULT 13  
US-09-710-279-3357/c  
; Sequence 3357, Application US/09710279  
; Patent No. 6703492  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/09/710,279  
; CURRENT FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3357  
; LENGTH: 4055  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-09-710-279-3357  
Alignment Scores:  
Pred. No.: 3,43e-24 Length: 4055  
Score: 307.50 Matches: 75  
Percent Similarity: 51.4% Conservative: 39  
Best Local Similarity: 33.8% Mismatches: 91  
Query Match: 12.6% Indels: 17  
DB: 3 Gaps: 8  
US-10-731-525-8 (1-466) x US-09-710-279-3357 (1-4055)  
QY 98 GluileMetAspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSer--- 116  
Db 2981 GAGATTTTAAATGGGCATATCAAACTATGAAATGATATGTTTATCTTCAGTTT 2922  
QY 117 GlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArgProPheArgVal 136  
Db 2921 GGTGCTGAGAGTATGCTTTTGTGATTGACTTCTCAAAATTAACCCGACGACAAAT 2862  
QY 137 PheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAspAlaVal--- 155  
Db 2861 GTATTTTATAGACACTGATTACATTTTCAGAAACATATGACTTAATAGTAGGGTAA 2802  
QY 156 GluLysHisTyrGlyIleArgIleGluTyrMetPheProasp---AlaValGluValGln 174  
Db 2801 GATAAATATCCGAATATACGATTAATAAATGAAAAACCCGAACTTACGTAGAGGAA 2742  
QY 175 AlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGluCys 194  
Db 2741 GGTGAGAAATATAATCCTGCTTTA-----TGAAGAATGATCCTAACCAATGTTGC 2691  
QY 195 ArgValArgLysValArgProLeuArgAlaLeuLysGlyLeuArgAlaTrpIleThr 214





; ORGANISM: Mycobacterium tuberculosis  
; OTHER INFORMATION: H37RV  
US-09-103-840A-1

## Alignment Scores:

Pred. No.:	3,37e-19	Length:	4411529
Score:	307.50	Matches:	98
Percent Similarity:	44.6%	Conservative:	38
Best Local Similarity:	32.1%	Mismatches:	124
Query Match:	3	Indels:	45
DB:	3	Gaps:	12

US-10-731-525-8 (1-466) x US-09-103-840A-1 (1-4411529)

Qy	34	ArgProIleGlyGlyAlaValAsnPheAsnLeuSerGlnArgArgSerLeuValLysPro	53
Db	2686275	CGACCG--GGTGGTGGCAACTTCGTCAACACCGCGAGGAGTGAACGCTTCGCGCA	2686331
Qy	54	ValAsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThrIleVal	73
Db	2686332	GTG---GGTCATCCGGCGCGAGGAGGACGCTCGCATGAGCGCGAGACACAC	2686382
Qy	74	AlaSerAlaSerGluThrLysGluGluAspPheGluGlnIle	89
Db	2686383	-----AGGTGACCGAACCGCAACTACTGAGCTGGCGCGCGAGCTGCC	2686430
Qy	90	AspLeuAspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsn	109
Db	2686431	GAATCTGACGCGCGCCACCGCACATGTTGGCTGGACCGAGAACCTTCGCG	2686487
Qy	110	AspIleAla	112
Db	2686488	GACATCGCGCGCGCGCGCGGTGAGCGGACATCGCGGTGACACGTCGCACTAC	2686547
Qy	113	IleAlaPheSerGlyAlaGluAspValAlaLeuIleGluThrAlaLysLeuThrGlyArg	132
Db	2686548	GTAAGTCTCTTCAACATGGCTGATGCGGTGCTGGTGTGGTGTGGCGCCCAAGGTGCGACCG	2686607
Qy	133	ProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPhe	152
Db	2686608	GGCGTACCGGTCTCTTCTTGATACCGGTGATCCACTTCGTCGACACATCGGCACGAGA	2686667
Qy	153	AspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAsp	171
Db	2686668	GATCGATCGAGTCCGCTATGACGTCGGGTGCTCAATGTCCTCGGAGCACACAGTG	2686727
Qy	172	GluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGln	191
Db	2686728	GCCGAGCAGGACGCACTGGGCAAGGACTTGTTCGCC-----CGCAACCCCAT---	2686778
Qy	192	GluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAla	211
Db	2686779	GAATGTCGCGGTGGCGCAGGTGCTTCCCTGGGCAAGACGCTGCTGTGCTACTCGCG	2686838
Qy	212	TrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValVal	231
Db	2686839	TGGGTGACCGGGTACGCGGGTGCATGACCG---ACCGGGCCAAATGCCCGCTGGTC	2686895
Qy	232	GlnValAspProAlaPheGluGlyMetAspGlyIleGlySerLeuValLysTrpAsn	251
Db	2686896	AGCTTCGATGAGACGTTTCAAA-----CTAGTGAAGGTCAAC	2686931
Qy	252	ProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValPro	271
Db	2686932	CCGCTGGCGCGGTGGACCGACCAAGATGTCCAGGAATACATTCGCGACACGACGTGCTG	2686991
Qy	272	ValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgPro	291
Db	2686992	GTTAATCCGCTTGGCGGAAGGCTATCCGTCGATGCTGCTGCTCGTCGACAGCCAA	2687051
Qy	292	ValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpGluAspAlaLys	311
Db	2687052	CCCGCGAAGGCGCGCACCGCGCAGCGGACG-----TGGCAGGGGCTGGCCAGACC	2687105

Qy	312	GluCysGlyLeuHis	316
Db	2687106	GAATGCGGTTGCAC	2687120

Search completed: February 20, 2006, 23:57:58  
Job time : 3412 secs

**THIS PAGE LEFT BLANK**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 20, 2006, 22:59:55 ; Search time 1395 Seconds  
(without alignments)  
2762.387 Million cell updates/sec

Title: US-10-731-525-8

Perfect score: 2445

Sequence: 1 MALAFSSISAPSTPSSSE.....KYPSEKRDVDSLMFVNALR 466

Scoring table:

BLOSUM62 Xgapop 10.0 , Xgapext 0.5  
Xgapop 10.0 , Ygapext 0.5  
Ygapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-O=/abs/ABSSWEB.spool/US10731525/runat.17022006.164909.374/app\_query.fasta\_1  
-DB=Published Applications\_NA\_Main -QFWT=fastap -SUFFIX=p2n.rnpbm  
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1  
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORES=pct  
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext  
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs06p  
-USER=US10731525 @CGN 1.1 1549 @runat.17022006.164909.374 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -JSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications\_NA\_Main.\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2445	100.0	1629	7	US-10-731-525-7
2	2445	100.0	1629	7	US-10-762-049-7
3	1951	79.8	1730	7	US-10-425-114-7777
4	1951	79.8	2774	7	US-10-424-599-134742
5	1923	78.7	1795	7	US-10-731-525-5
6	1923	78.7	1795	7	US-10-762-049-5
7	1811.5	74.1	1398	3	US-09-938-842A-498

8	1811.5	74.1	1398	3	US-09-938-842A-498	Sequence 498, App
9	1738	71.1	1365	3	US-09-938-842A-2305	Sequence 2305, Ap
10	1738	71.1	1365	3	US-09-938-842A-2305	Sequence 2305, Ap
11	1645.5	67.3	1827	7	US-10-731-525-9	Sequence 9, Appli
12	1645.5	67.3	1827	7	US-10-762-049-9	Sequence 9, Appli
13	1645.5	67.3	1840	8	US-10-739-930-2784	Sequence 2784, Ap
14	1639.5	67.1	1880	7	US-10-437-963-2751	Sequence 2751, A
15	1597.5	65.3	1380	7	US-10-169-667A-1	Sequence 1, Appli
16	1509.5	61.7	1553	8	US-10-425-115-83222	Sequence 83222, A
17	1444.5	59.1	1850	8	US-10-425-115-83219	Sequence 83219, A
18	1422.5	58.2	1927	8	US-10-425-115-83223	Sequence 83223, A
19	1295.5	53.0	1210	7	US-10-731-525-3	Sequence 3, Appli
20	1295.5	53.0	1210	7	US-10-762-049-3	Sequence 3, Appli
21	1262.5	51.6	1215	7	US-10-731-525-1	Sequence 1, Appli
22	1262.5	51.6	1215	7	US-10-762-049-1	Sequence 1, Appli
23	1241.5	50.8	1183	7	US-10-425-114-15300	Sequence 15300, A
24	1203.5	49.2	1156	7	US-10-425-114-12185	Sequence 12185, A
25	1105	45.2	711	7	US-10-424-599-135400	Sequence 135400, A
26	1039	42.5	1194	8	US-10-425-115-83221	Sequence 83221, A
27	999.5	40.9	1294	8	US-10-425-115-83224	Sequence 83224, A
28	914	37.4	549	8	US-10-425-115-147208	Sequence 147208, A
29	871.5	35.6	985	7	US-10-767-701-10835	Sequence 10835, A
30	754	30.8	796	7	US-10-424-599-82919	Sequence 82919, A
31	731.5	29.9	705	6	US-10-369-493-37482	Sequence 37482, A
32	716.5	29.3	471	3	US-09-732-627A-2057	Sequence 2057, Ap
33	697.5	28.5	585	7	US-10-021-323-3831	Sequence 3831, Ap
34	679	27.8	623	7	US-10-021-323-11704	Sequence 11704, A
35	571	23.4	742	8	US-10-425-115-12975	Sequence 12975, A
36	539	22.0	690	6	US-10-369-493-42894	Sequence 42894, A
37	537	22.0	330	7	US-10-767-701-17535	Sequence 17535, A
38	520	21.3	588	7	US-10-021-323-3903	Sequence 3903, Ap
39	509	20.8	552	7	US-10-021-323-11747	Sequence 11747, A
40	472.5	19.3	56609	9	US-10-915-740A-7	Sequence 7, Appli
41	472.5	19.3	2242716	9	US-10-915-740A-1068	Sequence 1068, Ap
42	454.5	18.6	599	7	US-10-424-599-69970	Sequence 69970, A
43	449.5	18.4	675	6	US-10-369-493-43366	Sequence 43366, A
44	439	18.0	514	3	US-09-770-961-142	Sequence 142, App
45	405.5	16.6	269223	7	US-10-672-787-41	Sequence 41, Appl

ALIGNMENTS

RESULT 1  
US-10-731-525-7  
; Sequence 7, Application US/10731525  
; Publication No. US20040121440A1  
; GENERAL INFORMATION:  
; APPLICANT: Falco, Saverio Carl  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167-C  
; CURRENT APPLICATION NUMBER: US/10/731,525  
; PRIOR FILING DATE: 2003-12-09  
; PRIOR APPLICATION NUMBER: US/09/720,318A  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 1998-07-14  
; NUMBER OF SEQ IDS: 12  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 7  
; LENGTH: 1629  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-10-731-525-7

Alignment Scores:			
Pred. No.:	3.84e-271	Length:	1629
Score:	2445.00	Matches:	466
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	7	Gaps:	0

```
US-10-731-525-8 (1-466) x US-10-731-525-7 (1-1629)
Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSerGlu 20
Db 32 ATGGCCCTCGCTTTCACCTCTCTCAATTTCCGACCAACTTCCACCTTCCCATCATCGGAA 91
Qy 21 ProLysLeuProGlnIleGlySerIleArgIleSerGluArgProIleGlyGlyAlaVal 40
Db 92 CCCAAATCTCCGCAAAATTTGGTCAATTTAGGATTTCCGAGAGGCCCATTTGAGGGCCCGTT 151
Qy 41 AsnPheAsnLeuSerGlnArgSerLeuValLysProValAsnAlaGluProProArg 60
Db 152 AATTTCAATTTATCTCAAGACCGGCTTGGTAAGCCCGTTAAGCCGGAACCTCCACGC 211
Qy 61 LysAspSerIleValProLeuAlaAlaThrThrIleValAlaSerAlaSerGluThrLys 80
Db 212 AAGGATTCATGTTCTCTCGACCAACACCATCGTTGCTTCTGCTTCTGAGACGAAA 271
Qy 81 GluGluAspPheGlnIleAlaSerAspLeuAspAsnAlaSerProLeuGluIleMet 100
Db 272 GAGGAAGATTTTGAACAGATAGCCAGTGATCTCGACCAATGCTTCACCTCTTGAATCATG 331
Qy 101 AspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAlaPheSerGlyAlaGluAsp 120
Db 332 GATAGAGCCCTCGACAAATTCGCGCAACGACATAGCTATTGCTTCAGTGGTGTGAAGAT 391
Qy 121 ValAlaLeuIleGluTyrAlaLysLeuThrGlyArgProPheArgValPheSerLeuAsp 140
Db 392 GTTGCCTTTGATTGATGATGCGAAATTCGCGGTGCGACCTTTTAGGGTTTTCAGTTTGGAC 451
Qy 141 ThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAspAlaValGluLysHisTyrGly 160
Db 452 ACTGGGAGACTGAACCCAGAAACTTATCAACTTTTGTGCGGTGAGAACATATTGGA 511
Qy 161 IleArgIleGluTyrMetPheProAspAlaValGluValGlnAlaLeuValArgSerLys 180
Db 512 ATTGCGATTGATACATGTTCCCTGATGCTGTTGAGGTTTCAGGCTTGGTGAGGTAAG 571
Qy 181 GlyLeuPheSerPheTyrGluAspGlyHisGlnGluCysArgValArgLysValArg 200
Db 572 GGGTATTCTCTTACGAGGATGGGCACCAAGAGTGTTCAGGGGTGAGAAAGGTGAGG 631
Qy 201 ProLeuArgAlaLeuLysGlyLeuArgAlaThrIleThrGlyGlnArgLysAspGln 220
Db 632 CCTTTAAGAGGGGCCCTTAAGGGTCTCAGACATGGATACTGTCTCAGAGGAAAGACGAG 691
Qy 221 SerProGlyThrArgSerGluIleProValValGlnValAspProAlaPheGluGlyMet 240
Db 692 TCACCTGGTACTAGGTCCTGAATACCGGTTGTTTTCAGGTTGATCGGCTTTTGAGGGAAATG 751
Qy 241 AspGlyGlyIleGlySerLeuValLysTrpAsnProValAlaAsnValLysGlyHisAsp 260
Db 752 GATGGTGGAAATGGAAGCTTGGTGAAGTGAACCTGTTGCAAAATGTGAAGGGCCATGAC 811
Qy 261 IleTrpAsnPheLeuArgThrMetAsnValProValAsnSerLeuHisAlaLysGlyTyr 280
Db 812 ATATGGAACTTCTTTAGGACCATGAATGTGCTGTGAATTCCTTGATGCAAAAGGATAT 871
Qy 281 ValSerIleGlyCysGluProCysThrArgProValLeuProGlyGlnHisGluArgGlu 300
Db 872 GTTTCCATTGGGTGTGAGCCCTGCACCTAGGCTGTGTTTACCTGGGCAACATGAAGGGAA 931
Qy 301 GlyArgTrpTrpTrpGluAspAlaLysAlaLysGluCysGlyLeuHisLysGlyAsnVal 320
Db 932 GGGAGGTGGTGGGAGGATGCCAAAGCTAAGGAATGTGGTCTCTTCACAAAAGGAAATGTA 991
Qy 321 LysGlnGlnLysGluGluAspValAsnGlyAsnGlyLeuSerGlnSerHisAlaAsnGly 340
Db 992 AAGCAGCAGAAAGAGGAGGTGTTAATGGAATGGGCTATCCCAATCCATGCAAAATGGT 1051
Qy 341 AspAlaThrThrValProAspIlePheAsnSerProAsnValValAsnLeuSerArgThr 360
Db 1052 GATGCTACCACTGTGCTGACATTTTCAACAGCCGGAATGTAGTTAACTTTCAGCAGGACT 1111

RESULT 2
US-10-762-049-7
; Sequence 7, Application US/10762049
; Publication No. US20040139492A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/10762,049
; PRIOR FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: US/09/720,318A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 1629
; TYPE: DNA
; ORGANISM: Glycine max
US-10-762-049-7

Alignment Scores:
Pred. No.: 3,84e-271 Length: 1629
Score: 2445.00 Matches: 466
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 7 Gaps: 0

US-10-731-525-8 (1-466) x US-10-762-049-7 (1-1629)
Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSerGlu 20
Db 32 ATGGCCCTCGCTTTCACCTCTCTCAATTTCCGACCAACTTCCACCTTCCCATCATCGGAA 91
Qy 21 ProLysLeuProGlnIleGlySerIleArgIleSerGluArgProIleGlyGlyAlaVal 40
Db 92 CCCAAATCTCCGCAAAATTTGGTCAATTTAGGATTTCCGAGAGGCCCATTTGAGGGCCCGTT 151
Qy 41 AsnPheAsnLeuSerGlnArgSerLeuValLysProValAsnAlaGluProProArg 60
Db 152 AATTTCAATTTATCTCAAGACCGGCTTGGTAAGCCCGTTAAGCCGGAACCTCCACGC 211
Qy 61 LysAspSerIleValProLeuAlaAlaThrThrIleValAlaSerAlaSerGluThrLys 80
Db 212 AAGGATTCATGTTCTCTCGACCAACACCATCGTTGCTTCTGCTTCTGAGACGAAA 271
```



```
Db 457 GGTGCGACCTACAGAGTGTAGTCTTGACACTGGGAGACTGAATCCAGAAACCTACAAA 516
Qy 151 LeupheaspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProaspAla 170
Db 517 TTTTTCGATCGGTGAGAGCATATATGGAATTCACATTGAGTACATGTTCCCTGATGCG 576
Qy 171 ValGluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHis 190
Db 577 GTTGAGGTTCAGGCATTTAGTAAGCACTAAAGGCTCTTCATTTTACGAGATGGGCAT 636
Qy 191 GlnGluCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArg 210
Db 637 CAAGAGTGCTGCCGAGTAAGAGGTGAGGCCCTTGAGGAGAGCCCTTAAGGGTCTCAAA 696
Qy 211 AlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProVal 230
Db 697 GCATGGATCATCTGACACAGAAAGATCAGTCTCTCGTACTAGTCTGAAATCCCTGTT 756
Qy 231 ValGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrp 250
Db 757 GTCCAGGTTGATCTGTTTGTAGGGACTGGATGGTGGAAATGGCAGCCCTGGTGAAGTGG 816
Qy 251 AsnProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgTrpMetAsnVal 270
Db 817 AACCCGGTTCGAATGTTATGGTCTAGACATATGGAGCTTCCTTAGGACCATGGATGTT 876
Qy 271 ProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArg 290
Db 877 CCTGTAAATTCATTGTCATCTCCCAAGGATATGTTTCGATTTGGCTGTGAGCCGTGCACAAG 936
Qy 291 ProValLeuProGlyGlnHisGluArgGluGlyValArgTrpTrpGluAspAlaLysAla 310
Db 937 CCAGTTTTACCGGTCAACATGAAGAAGAGAAAGGTGGTGGTGGAGAGATGCCAAAGCC 996
Qy 311 LysGluCysGlyLeuHisLysGlyAsnValLysGlnLysGluGluAspValAsnGly 330
Db 997 AAGGAGTGTGCTTCACAAAGGTAATATAAAACATGAAGATGCTGCCAGCTTAATGGA 1056
Qy 331 AsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrValProAspIlePheAsn 350
Db 1057 AATGGGCGCTCCCAA-----GCAAAATGGATCTGCC---ACTGTCGCTGACATTTTCAAC 1107
Qy 351 SerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAsp 370
Db 1108 TCCAGAGATGGTCACTTGACAGAGTCTGGGATTCAGAAATTTGGCAAAATTCGAGAAC 1167
Qy 371 ArgLysGluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMet 390
Db 1168 CGAAAGAACCTCGCTTGTGTCTCTATGCACCATGGTGTGCGCTTCTGTGAGGCTATG 1227
Qy 391 GluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGly 410
Db 1228 GAGGAATCATATGTTGATTTGGCAAAAGTTAGCTGGGTCA---GGAGTGAAGGTTGCA 1284
Qy 411 LysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySer 430
Db 1285 AAATTCAGCGGATGAGACCAAGGAATATCAAGACTGAACCTGACGATTCGAGTGGGAGC 1344
Qy 431 PheProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSer 450
Db 1345 TTCCCCACAATFACTCTCTCCCCCAACACCTCTTCTCAACCA---ATTAAGTACCCTTGG 1401
Qy 451 GluLysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
Db 1402 GAAAGAGAGATGTTGATTCATTGACGCGCATTCGTGAATGCTTTACCG 1449
```

## RESULT 4

US-10-424-599-134742  
; Sequence 134742, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21 (53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 134742  
; LENGTH: 2774  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_92682C.1  
US-10-424-599-134742

Alignment Scores: 8.53e-214 Length: 2774  
Pred. No.: 1951.00 Matches: 382  
Score: 88.2% Conservative: 38  
Best Local Similarity: 80.3% Mismatches: 40  
Query Match: 79.8% Indels: 16  
DB: 7 Gaps: 7

US-10-731-525-8 (1-466) x US-10-424-599-134742 (1-2774)

```
Qy 1 MetAlaLeuAlaPhe-----ThrSerSerIleSerAlaProThrSerThr 15
Db 106 ATGGCTCTCGCCGTTTCCACTACTTCTCTTCAGCAGCTGCTGCAGCTTCTAGCTCC 165
Qy 16 Phe-----ProSerSerGluProLysLeuProGlnIleGlySerIleArg 30
Db 166 TTCTTCTCTCGTCTCGGATCTTCTCGACGCAAAAGCTCTCAAAATCGTTCGTTTCG 225
Qy 31 IleSerGluArgProIleGlyAlaValAsnPheAsnLeuSerGlnArgSerIleu 50
Db 226 TTTCCGAGAGGTGCGTCTGTTCTGTTGTGTCATTAATGTAACCAACGCGCTCCTTG 285
Qy 51 ValLysProValAsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThr 70
Db 286 GTGAGGCCACTCAACGCCGAAACCGCAACGGAAGATTCTATTGTTCTCTTTCAGCA--- 342
Qy 71 ThrIleValAlaSerAlaSerGluThrLysGluGluAspPheGluGlnIleAlaSerAsp 90
Db 343 ACTATCGTTCCTCTGAGGTTGAGAGGAAGAGAGAGATTTCAGCAAAATAGCGAAAGAC 402
Qy 91 LeuAspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAsp 110
Db 403 CTTGAAAATGCAATCTCCTCTTGAATAATTATGATAGGCGCTCGAGAAAATTTGGGAACGAC 462
Qy 111 IleAlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThr 130
Db 463 ATCGCTATTGCATTTAGTGTGCTGGAAGATGTTCTTTGATTGAGTATGCACATTTGACG 522
Qy 131 GlyArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGln 150
Db 523 GGTGCACTTACAGAGTGTGTTAGTCTTGACACTGGGAGACTGAATCCAGAAACCTACAA 582
Qy 151 LeuPheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAla 170
Db 583 TTTTTCGATCGGTGAGAGCATTTATGGAATTCACATTGAGTACATGTTCCCTGATGCG 642
Qy 171 ValGluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHis 190
Db 643 GTTCAGGTTCAGGCATTTAGTAAGCACTAAAGGCTCTTCTCATTTTACGAGATGGGCAT 702
Qy 191 GlnGluCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArg 210
Db 703 CAAGAGTGCTGCCGAGTAAGAGGTGAGGCCCTTGAGGAGAGCCCTTAAGGGTCTCAAA 762
Qy 211 AlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProVal 230
Db 763 GCATGGATCACTGGACAGAGAAAAGATCAGTCTCCTCGTACTAGTCTGAAAATCCCTGTT 822
```



```
Qy 231 ValGlnValAspProAlaPheGluGluMetAspGlyGlyIleGlySerLeuValLysTrp 250
Db 823 GTCCAGGTGATCTGTTTGGGGAGCTGGATGGTGAATGGCAGCCTGGTGAAGTGG 882
Qy 251 AsnProValAlaAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgTrpMetAsnVal 270
Db 883 AACCCGGTTCGAATGTAATGGTCTAGACATATGGAGCTTCCTTAGGACCATGGATGTT 942
Qy 271 ProValAsnSerLeuHisAlaLysGlyValSerIleGlyCysGluProCysThrArg 290
Db 943 CTTGTAAATTCATTTGCATTCCTCAAGGATATGTTTCGATTTGGCTGTGAGCCGTGCACAAAG 1002
Qy 291 ProValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpGluAspAlaLysAla 310
Db 1003 CCAGTTCCTACCGGTCAACATGAAGAGAGAGAGAGTGGTGGTGGAGGATGCCAAAGCC 1062
Qy 311 LysGluCysGlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluAspValAsnGly 330
Db 1063 AAGGAGTGTGGTCTTCAACAAGGTAATATATAAATCATGAAGATGCTGCCAGCTTAATGGA 1122
Qy 331 AsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsn 350
Db 1123 AATGGGGCTCCCAA-----GCNAATGGATCTGCC---ACTGTGCTGACATTTTCAAC 1173
Qy 351 SerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAsp 370
Db 1174 TCCAGGATGTGGTCAGCTTTCAGCAGGTCTGGGATTCAGAAATTTGGCAAAATTTGGAGAAC 1233
Qy 371 ArgLysGluProTrpLeuValValLeuTrpAlaProTrpCysProTrpCysGlnAlaMet 390
Db 1234 CGAAAGAACCTGGCTGTGTGCTCTATGCACCATGGTGGCGCTTCTGTCAAGGCTATG 1293
Qy 391 GluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGly 410
Db 1294 GAGGAATCATATGTTGATTTGGCNAAGATAGCTGGTCA---GGAGTGAAGGTGGCA 1350
Qy 411 LysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySer 430
Db 1351 AAATTCAGAGCGATGAGACAGAGAGAGATATGCAAAAGCTGAACCTGCAGTGGGAAGC 1410
Qy 431 PheProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTrpProSer 450
Db 1411 TTCCCAACAATACTACTCTTCCCAACACACTCTTCTCAACCA---ATTAAGTACCCCTTCG 1467
Qy 451 GluLysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
Db 1468 GAAAGAGAGATGTTGATTCATTGACGGCAATTCGTGAATGCTTTACGG 1515
```

## RESULT 5

```
US-10-731-525-5
; Sequence 5, Application US/10731525
; Publication No. US20040121440A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/10/731,525
; PRIOR FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: US/09/720,318A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1795
; TYPE: DNA
; ORGANISM: Glycine max
US-10-731-525-5
Alignment Scores:
```

```
Pred. No.: 7,64e-211 Length: 1795
Score: 1923.00 Matches: 378
Percent Similarity: 87.2% Conservative: 37
Best Local Similarity: 79.4% Mismatches: 45
Query Match: 78.7% Indels: 16
DB: Gaps: 7

US-10-731-525-8 (1-466) x US-10-731-525-5 (1-1795)

Qy 1 MetAlaLeuAlaPheThrSerSerIleSer-----AlaProThrSerThr 15
Db 56 CTTGCCGTTTCCACTACTTCTCTCTAGCTGCAGCAGCAGCAGCGTGCAGCTCT 115
Qy 16 Phe-----ProSerSerGluProLysLeuProGlnIleGlySerIleArg 30
Db 116 TTCTTCTCGCGCTTGGATCTTCATCGAGCTAAAGCTCCGCAAAATTTGTTCTCTTCGG 175
Qy 31 IleSerGluArgProIleGlyGlyAlaValAsnPheAsnLeuSerGlnArgSerIleu 50
Db 176 TTTCCGAGAGAGGCTCAAGTTTCGTCGTGTTGTTAAATTTAACTCAAGACGCTCCTCG 235
Qy 51 ValLysProValAsnAlaGluProAspLysAspSerIleValProLeuAlaAlaThr 70
Db 236 GTGAGGCTCACTCAATGCCAACCCGCAACGGAATGTTCTGTGTTCTCTCTTCAGCA--- 292
Qy 71 ThrIleValAlaSerAlaSerGluThrLysGluGluAspPheGluGlnIleAlaSerAsp 90
Db 293 ACTATCGTCTCTCGAGTTGAGAGGAGAGAGAGAGATTTTTCAGCAATTTAGCGAAGAC 352
Qy 91 LeuAspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAsp 110
Db 353 CTTGAAAATTTCTCTCTCTGAGATTTATGGAATTTAGGACCTTCGAGAAATTTGGGAACGAC 412
Qy 111 IleAlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThr 130
Db 413 ATCGCTATTGCTCTTGTGCTGCTGAAGATGTTGCTTTGATTGAGTATGCACATTTTCACG 472
Qy 131 GlyArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGln 150
Db 473 GGTGCGACCTTACAGAGTGTGTTAGTCTTGACACTGGGAGACTGAAACCCAGAAACCTACAA 532
Qy 151 LeuPheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAla 170
Db 533 TTTTTCAGCGCTCTGAGAGCAATTTATGGAATTTATGATTCATATTCAGTACATGTTCCCTGATGCG 592
Qy 171 ValGluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHis 190
Db 593 GTTGAGGTTTACGAGCATTTAGTAAGAACTAAGGGGCTCTTCTCATTTTACGAGGATGGGCAT 652
Qy 191 GlnGluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArg 210
Db 653 CAAGAGTGTGTAGAGTAAAGAGGTGAGGCCCTTGAGGAGAGCCCTTAAGGGTCTCAAA 712
Qy 211 AlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProVal 230
Db 713 GCATGGAATTTACTGGACAGAGAAAGACCACTCTCTGGTACTAGGTCTGAAATCCCTATT 772
Qy 231 ValGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrp 250
Db 773 GTCAGGTTGATCCTGTTTGTGAGGACTGGATGGTGGAAATTTGGCAGCCTGGGAAAGTGG 832
Qy 251 AsnProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnVal 270
Db 833 AACCCGGTTCGAATGTTAATGCTTAGACATATGGAATCTCTTAGGACCATGAATGTT 892
Qy 271 ProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArg 290
Db 893 CTTGTAATTTCTATGCAATCCCAAGGATATGTTTCGATTTGGCTGTGAGCCATGCACAAGG 952
Qy 291 ProValLeuProGlyGlnHisGluArgGluGlyArgTrpTrpGluAspAlaLysAla 310
Db 953 CCGGTTTATACCCGGAACAATGATGAAGAGAGAGAGTGGTGGTGGGAGGATGCGCAAGCC 1012
```

```
Qy 311 LysGluCysGlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluAspValAsnGly 330
Db 1013 AAGAGTGTGGTCTTCAAAAGGTAATTTGAAACAGAGAGATGCTGCCAGCTTAATGGA 1072
Qy 331 AsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsn 350
Db 1073 AATGGGACCTCCCAA-----GGAAATGGCTCTGCC--ACTGTTGCTGCATTTTCATC 1123
Qy 351 SerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAsp 370
Db 1124 TCCAGAAATGGTTCAGCTTGAGCAGGTCCGGGATTGAGAAATTTGGCAAAATTAGAGAAC 1183
Qy 371 ArgLysGluProTropLeuValLeuTyrAlaProTropCysProTyrCysGlnAlaMet 390
Db 1184 CGAAAGAACACTGGCTTGTTGCTCTATGCAACATGGTGGCCCTCTCTGTCAGGCTATG 1243
Qy 391 GluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGly 410
Db 1244 GAGGAGTCGTATGTTGATCTGGCAGAGAGATTAGCAAGGTCA--GGAGTGAAGGTTGCA 1300
Qy 411 LysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLysGlySer 430
Db 1301 AAATTCAGAGCGGATGGAGAGCAGAGAGAAATATGCAAAAGAGTGAACCTGCAGTTGGGAAGC 1360
Qy 431 PheProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSer 450
Db 1361 TTCCCCAACAACTCTCTCTCCCAAGACCTCTTCTCAACCA---ATTAGTACCCCTCA 1417
Qy 451 GluLysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
Db 1418 GAAAGAGAGATGTTGATTCATTGACGGCATTCGTGAATGCTTAACG 1465

RESULT 6
US-10-762-049-5
; Sequence 5, Application US/10762049
; Publication No. US20040139492A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-C
; CURRENT APPLICATION NUMBER: US/10762,049
; CURRENT FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: US/09/720,318A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1795
; TYPE: DNA
; ORGANISM: Glycine max
US-10-762-049-5

Alignment Scores:
Pred. No.: 7,64e-211 Length: 1795
Score: 1923.00 Matches: 378
Percent Similarity: 87.2% Conservative: 37
Best Local Similarity: 79.4% Mismatches: 45
Query Match: 78.7% Indels: 16
DB: Gaps: 7

US-10-731-525-8 (1-466) x US-10-762-049-5 (1-1795)
Qy 1 MetAlaLeuAlaPheThrSerSerIleSer-----AlaProThrSerThr 15
Db 56 CTGGCGTTTCCACTACTTCTCTCTCTCAGCTGCAGCAGCAGCAGCGCTCGAGCTCT 115
Qy 16 Phe-----ProSerSerGluProLysLeuProGlnIleGlySerIleArg 30
Db 116 TTCTTCTCGCGCTTGGATCTTCTATCGACGCTAAAGCTCCGCAATTTGGTTCCTTCGG 175
```

```
Qy 31 IleSerGluArgProIleGlyGlyAlaValAsnPheAsnLeuSerGlnArgSerLeu 50
Db 176 TTTCGGGAGAGGCTCAAGTTTCGTCTGGTGTGTTAAATTTAACTCAAGACGCTCCTCG 235
Qy 51 ValLysProValAsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThr 70
Db 236 GTGAGGCCACTCAATCGCAACCGCAACGGAATGATTCTGTGTCTCTCTTGCAGCA-- 292
Qy 71 ThrIleValAlaSerAlaSerGluThrLysGluGluAspPheGluGlnIleAlaSerAsp 90
Db 293 ACTATCGTTCTCTCCTGAGGTGAGAGGAGGAGAGAGATTTTGGACAAATTAGCAAGAC 352
Qy 91 LeuAspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAsp 110
Db 353 CTTGAAAATTCATCTCTCTTGAGATTATGATAAGGCCCTCGAGAAATTTGGGAACGAC 412
Qy 111 IleAlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThr 130
Db 413 ATCGCTATTGCCCTTAGTGTGCTGAAGATGTTGCTTGTGATTGAGTATGCACATTTGACG 472
Qy 131 GlyArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrGln 150
Db 473 GGTGCACTTACAGAGTGTGTTAGTCTTGACACTGGGAGACTGAACCCAGAAACCTACAA 532
Qy 151 LeuPheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAla 170
Db 533 TTTTGTGACGCTGTTGAGAGCATTTATGGAATTCATATTGAGTACATGTTCCCTGATGCG 592
Qy 171 ValGluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHis 190
Db 593 GTTGAGGTTGAGGCATTTAGTAAGAACTTAAGGGGCTCTCTCTCATTTTACGAGGATGGGCAT 652
Qy 191 GlnGluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArg 210
Db 653 CAAGAGTGTGTAGAGTAAGAAAGGTGAGGCCCTTGAGGAGAGCCCTTAAGGGTCTCAAA 712
Qy 211 AlaTrrIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProVal 230
Db 713 GCATGGNTTACTGGACAGAGAAAGACCACTCTCTGTTACTAGTCTCGAAATCCCTATT 772
Qy 231 ValGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrrp 250
Db 773 GTCCAGGTTGATCCTCTGTTTGTGAGGACTGGATGGTGGAAATTTGGCAGCCTGGTGAAGTGG 832
Qy 251 AsnProValAlaAsnValLysGlyHisAspIleTrrpAsnPheLeuArgThrMetAsnVal 270
Db 833 AACCCGGTTGCAATGTTAATGGTCTAGACATATGGAACCTTCCTTAGACCATGAATGTT 892
Qy 271 ProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArg 290
Db 893 CTTGTAATTCATTGCAATTCCTCAAGGATATGTTTCGATTGGCTGTGAGCCATGCACAAGG 952
Qy 291 ProValLeuProGlyGlnHisGluArgGluGlyArgTrrpTrrpGluAspAlaLysAla 310
Db 953 CCGGTTTTACCGGCAACATGAAAGAGAAAGGTTGGTGGTGGGAGATGCCAAAGGCC 1012
Qy 311 LysGluCysGlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluAspValAsnGly 330
Db 1013 AAGGAGTGTGCTCTTCAAAAGGTAATTTGAAACAGAGAGATGCTGCCAGCTTAATGGA 1072
Qy 331 AsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsn 350
Db 1073 AATGGGACCTCCCAA-----GGAAATGGCTCTGCC--ACTGTTGCTGCATTTTCATC 1123
Qy 351 SerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAsp 370
Db 1124 TCCAGAAATGGTTCAGCTTGAGCAGGTCCGGGATTGAGAAATTTGGCAAAATTAGAGAAC 1183
Qy 371 ArgLysGluProTrrpLeuValValLeuTyrAlaProTrrpCysProTyrCysGlnAlaMet 390
Db 1184 CGAAAGAACAACTGGCTTGTGCTCTATGCAACATGGTGGCCGCTCTGTGAGGCTATG 1243
Qy 391 GluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGly 410
```

```
Db 1244 GAGGAGCGTATGTTGATCTGGCAGAGAGTTAGCAAGGTCA---GGAGTGAAGGTTGCA 1300
Qy 411 LysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySer 430
Db 1301 AATTCAGACCGATGAGAGCAGAGCAATATGCAAAAGTGAATCTGCAGTTGGGAAGC 1360
Qy 431 PheProThrIleLeuPhePheProLysHisSerSetArgProThrIleLysTyrProSer 450
Db 1361 TTCCCAACAATCTCTCTCTCCCAAGCACTCTCTCAACCA---ATTAAGTACCCCTTCA 1417
Qy 451 GluLysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
Db 1418 GAAAGAGAGATGTTGATTCATTCAGCGCATTCGTGAATGCCTTACGG 1465
```

## RESULT 7

```
US-09-938-842A-498
; Sequence 498, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 498
; LENGTH: 1398
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-498
```

```
Alignment Scores:
Pred. No.: 3,89e-198 Length: 1398
Score: 1811.50 Matches: 351
Percent Similarity: 84.0% Conservative: 48
Best Local Similarity: 73.9% Mismatches: 57
Query Match: 74.1% Indels: 19
DB: 3 Gaps: 7
```

US-10-731-525-8 (1-466) x US-09-938-842A-498 (1-1398)

```
Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSer--- 19
Db 1 ATGGCAATGCTGTAAATGTTCTCTCTCTGCTCTCGGATCATATAAATCTCGTTTC 60
Qy 20 -----GluProLysLeuProGlnIleGlySerIleArgIleSerGluArgPro 35
Db 61 GGTGTTTCATGGAGCCAAAGTTTCGCAATTTGGTTCGTTAGGTATTGGATCGTGT 120
Qy 36 IleGlyGlyAlaValAsnPheAsnLeuSer---GlnArgArgSerLeuValLysProVal 54
Db 121 CATGTTGCTCCTGTCTCTGTAATCTATCTGGGAAGCGATCATCTGTGTTAAACCTTTA 180
Qy 55 AsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThrThrIleValAla 74
Db 181 AACGCTGAACCAAGACAAAGAGATTCAATGATTCTCTTCGCGCAACAATGGTAGCAGAA 240
Qy 75 SerAlaSerGluThrLysGlu-----GluAspPheGluGlnIleAlaSerAspLeu 91
Db 241 ATTGCAGAGAGAGTTGAGTGGTGGATTGAGATTGAGATTGAGAGCTTGTAGAGATTGA 300
Qy 92 AspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIle 111
```

```
Db 301 GAGAAATGCTTACCTCTTGAGATTATGCAAAAGCTCTTGAGAAATACATCGGAAACGATATC 360
Qy 112 AlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGlyTyrAlaLysLeuThrGly 131
Db 361 GCCATTGCAATTAGTGGTGCAGAGATGTTGCTCTTATTAGTACGCTCATTTGACTGGG 420
Qy 132 ArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeu 151
Db 421 AGCCATTTAGAGTATTTAGTTTGGATACAGGAGGTTGAATCCTGAGACGTATCGGTTT 480
Qy 152 PheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaVal 171
Db 481 TTCGATGCGGTGGAGAACCACTATGGGATTAGAGATTGAGTATATGTTTCTGATTCTGTT 540
Qy 172 GluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGln 191
Db 541 GAGGTTCAAGGTTTGGTTAGGACCAAGGATGTTCTCTCTTTTATAGGATGGTCAATCAG 600
Qy 192 GluCysCysArgValArgLysValArgProLeuArgAlaLeuLysGlyLeuArgAla 211
Db 601 GAGTGTTCGCGTCTCGAAAGGTGAGACCTTTCAGGCGTCTCTCAAGGGTTTAAAGGCT 660
Qy 212 TrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValVal 231
Db 661 TGGATTACTGCTCAGAGGAAGATCAATCTCCGGGACACAGGCTCTGAGATTCCGGTGT 720
Qy 232 GlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsn 251
Db 721 CAGGTTGATCCGGTGTGAAGTTTGGATGGTGGAGTTGTTAGTTGGTGAAGTGGAAAT 780
Qy 252 ProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValPro 271
Db 781 CCGGTTTGGCAATGTTCAAGGGAATGATGTTTGGAACTTCTTGAGGACTATGGATGTTCCG 840
Qy 272 ValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgPro 291
Db 841 GTTAAACACATTCGATCGGCAGGATATATATCGATTGGATGTGAGCCTTCACGAAAGCG 900
Qy 292 ValLeuProGlyGlnHisGluArgGluGlyArgTyrTrpTrpGluAspAlaLysAlaLys 311
Db 901 GTTTTACCGGTCAGCACAGAGAGAGGAGATGTTGGTGGGAAGATGCTTAAGCCAG 960
Qy 312 GluCysGlyLeuHisLysGlyAsnValLysGlnLysGluLeuAspValAsnGlyAsn 331
Db 961 GAATGTGGACTTCACAAAGGGAATGTCAAA---GAAAACTCCGATGATGCTAAAGTGAAC 1017
Qy 332 GlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSer 351
Db 1018 GGGGAATCGAAATCC-----GCTGTTGCAGATATCTTTAAGAGT 1056
Qy 352 ProAsnValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArg 371
Db 1057 GAGAAATCTGTGACTTTGACGAGCGAGGAGATTGAGAAATTTGATGAAGTTGGAGAACCGT 1116
Qy 372 LysGluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMetGlu 391
Db 1117 AAAGAGCCTTGGATCGTGTGCTTTATGCTCGGTGGTGGCCCTTTTGTCAAGCCATGGAA 1176
Qy 392 GluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLys 411
Db 1177 GCATCGTATGATGACGCGGATAAATTTGGCTGGGAAGT---GGGATTAAAGTTGGCCAA 1233
Qy 412 PheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPhe 431
Db 1234 TTCAGAGCAGATGGTCAAGAGGAGTTTCTTAACAGGAATTCAGCTCGGTAGCTTC 1293
Qy 432 ProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSerGlu 451
Db 1294 CCTACCATCTGTTTTTCCCTTAAGAACTCATCGAGACCG---ATCAAGTATCGTCTGAG 1350
Qy 452 LysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
```

Db 1351 AAGAGAGATGTTGAGTCTTTGACTTCTGCTTCTTGAATCTTGTCCGA 1395

RESULT 8

US-09-938-842A-498

Sequence 498, Application US/09938842A

Publication No. US20040009476A9

GENERAL INFORMATION:

APPLICANT: Harper, Jeff

APPLICANT: Kreps, Joel

APPLICANT: Wang, Xun

APPLICANT: Zhu, Tong

TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

TITLE OF INVENTION: SAME, AND METHODS OF USE

FILE REFERENCE: SCRIPI300-3

CURRENT FILING DATE: 2001-08-24

PRIOR FILING DATE: 2001-08-24

PRIOR FILING DATE: 2000-08-24

PRIOR APPLICATION NUMBER: US 60/264,647

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/300,111

PRIOR FILING DATE: 2001-06-22

NUMBER OF SEQ ID NOS: 5379

SEQ ID NO 498

LENGTH: 1398

TYPE: DNA

ORGANISM: Arabidopsis thaliana

US-09-938-842A-498

Alignment Scores:

Pred. No.:	3.89e-198	Length:	1398
Score:	1811.50	Matches:	351
Percent Similarity:	84.0%	Conservative:	48
Best Local Similarity:	73.9%	Mismatches:	57
Query Match:	74.1%	Indels:	19
DB:	3	Gaps:	7

US-10-731-525-8 (1-466) x US-09-938-842A-498 (1-1398)

Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSer--- 19

Db 1 ATGGCAATGCTGTAATATTTCTTCTTCTCGTCTTCGGATCATATAAATCTCGTTTC 60

Qy 20 -----GluProLysLeuProGlnIleGlySerIleArgIleSerGluArgPro 35

Db 61 GGTGTTTCATTTGGAGCCAAAAGTTTCGCAATATGTTGTTGAGGTTATTGGATCGTGT 120

Qy 36 IleGlyGlyAlaValAsnPheAsnLeuSer---GlnArgArgSerLeuValLysProVal 54

Db 121 CATGTTGCTCTGTGCTCTGAATCTATCTGGGAGCGATCATCTGTTTAAACCTTTA 180

Qy 55 AsnAlaGluProProArgLysAspSerIleValProLeuAlaAlaThrIleValAla 74

Db 181 AACGCTGAACCAAGACAAAGATTCATGATTTCTTCTTGGCAACAATGGTAGAGAA 240

Qy 75 SerAlaSerGluThrLysGlu-----GluAspPheGluGlnIleAlaSerAspLeu 91

Db 241 ATTGCAGAGGAAGTTGAAGTGTGAGATTGAGGATTTTGAAGAGCTTTGCTAAGAAGTTA 300

Qy 92 AspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIle 111

Db 301 GAGAAATGTTTACCTCTTGAGATTATGACAAAGCTCTTGAGAAATACGGGAACATATC 360

Qy 112 AlaIleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGly 131

Db 361 GCCATTGCAATTTAGTGGTGCAGAGATGTTGCTCTTATTGAGTACGCTCATTTGACG 420

Qy 132 ArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeu 151

Db 421 AGGCCATTTAGATATTTAGTTTGGATACAGGAGGTTGAATCTCTGAGACGATTCGGT 480

Qy 152 PheAspAlaValGluLysHieTyrGlyIleArgIleGluTyrMetPheProAspAlaVal 171

Db 481 TTCGATGCGGTGGAGAAGCACATATGGATTAGGATTAGTATATGTTTCTCGATTCTGTT 540

Qy 172 GluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGln 191

Db 541 GAGGTTCAAGGTTTGGTTAGGACCAAGGATTTCTTCTTTTATGAGGATGGTCAACG 600

Qy 192 GluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLysValAla 211

Db 601 GAGTGTTCGCGTGTTCGAAAGGTGAGACCTTTGAGCGTGCTCTCAAGGTTTAAAGCT 660

Qy 212 TrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValVal 231

Db 661 TGGATTACTGTCAGAGGAAGATCAATCTCCGGGACAAGTCTGAGATTCCGGTTGTT 720

Qy 232 GlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsn 251

Db 721 CAGGTTGATCCGTTGTTGAAGGTTTGGATGGTGGAGTTTGTAGTTTGGTGAAGTGAAT 780

Qy 252 ProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValPro 271

Db 781 CCGGTTGCGAATGTTGAAGGAATGATGTTTGAACCTCTTGAGACTATGGATGTTCCG 840

Qy 272 ValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgPro 291

Db 841 GTTAACACATTCGATCGCGCAGGATATATATCGATTGATGTCGATGTCACGAAAGCG 900

Qy 292 ValLeuProGlyGlnHisGluArgGluGlyArgTyrTrpTrpGluAspAlaLysAlaLys 311

Db 901 GTTTTACCGGTCAGCACGAGAGAGAGAGAGTGGTGGTGGGAAGATGCTTAAAGCCAA 960

Qy 312 GluCysGlyLeuHisLysGlyAsnValLysGlnLysGluGluAspValAsnGlyAsn 331

Db 961 GAATGTGACTTACAAAGGAATGTCTCAA---GAAACTCCGATGATGCTAAAGTGAAC 1017

Qy 332 GlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSer 351

Db 1018 GGGGAATCGAAATCC-----GCTGTTGCAGATATCTTTAAGAGT 1056

Qy 352 ProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArg 371

Db 1057 GAGAATCTTGATCTTGGAGCAGCAGGAGATTGAGAATTTGATGAAGTTGGAGAACCGT 1116

Qy 372 LysGluProThrLeuValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMetGlu 391

Db 1117 AAAGAGCTTGGATGCTGCTGCTTATGCTCCGCTGCTGCTGCTTCTTCAAGCATGGAA 1176

Qy 392 GluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLys 411

Db 1177 GCATCGTATGATGAATCGCGGATAAATTGGCTGGAAGT---GGGATTAGGTTGCCAAA 1233

Qy 412 PheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPhe 431

Db 1234 TTCAGACAGATGTTGTCACCAAGAGAGTTCCTAAGCAGGAATTCGAGCTCGGTAGCTTC 1293

Qy 432 ProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTrpSerGlu 451

Db 1294 CCTACCATTCCTGTTTCCCTTAAGACTCATTCGAGACCG---ATCAAGTATCCGCTCAG 1350

Qy 452 LysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466

Db 1351 AAGAGAGATGTTGAGTCTTTGACTTCTTGAATCTTGTCCGA 1395

## RESULT 9

US-09-938-842A-2305

Sequence 2305, Application US/09938842A

Patent No. US20020160378A1

GENERAL INFORMATION:

APPLICANT: Harper, Jeff

APPLICANT: Kreps, Joel

APPLICANT: Wang, Xun

APPLICANT: Zhu, Tong

TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

TITLE OF INVENTION: SAME, AND METHODS OF USE

```
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2305
; LENGTH: 1365
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-938-842A-2305

Alignment Scores:
Pred. No.: 1,14e-189 Length: 1365
Score: 1738.00 Matches: 346
Percent Similarity: 82.5% Conservative: 45
Best Local Similarity: 73.0% Mismatches: 55
Query Match: 71.1% Indels: 28
DB: 3 Gaps: 9

US-10-731-525-8 (1-466) x US-09-938-842A-2305 (1-1365)

QY 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThr---SerThrPhe----- 16
DB 1 ATGGCTTTAGCTGTTACTTCTTCTCACTGCAATCTCTGGATCGAGTTTCTCTCGTTCT 60
QY 17 ---ProSerSerGluProLysLeuProGlnIleGlySerIleArgIleSerGluArgPro 35
DB 61 GGAGCTTCTTCTGAATCTAAAGCTCTTCAATATGTTGATAGTTATCTGATCGAACC 120
QY 36 IleGlyGlyAlaValAsnPheAsnLeuSerGlnArgArgSerLeuValLysProValAsn 55
DB 121 -----CAITGTCTCAGAGAGCTTACTCTATGAAACCTTTAAAC 159
QY 56 AlaGluPro---ProArgLysAspSerIleValProLeuAlaAlaThrThrIleValAla 74
DB 160 GCTGAGTCACATTCACGACGGAATCTTGGGTACTCGTTCACCTCAATTTGTCCT 219
QY 75 SerAlaSerGluThr-----LysGluGluAspPheGluGlnIleAlaSerAspLeuAsp 92
DB 220 GAAGTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 279
QY 93 AsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAla 112
DB 280 GATGCTTCTCCACTTGAAATCATGGATAAAGCTCTTGAGAGATTCGGAGACCAATCGCA 339
QY 113 IleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArg 132
DB 340 ATTGCTTTTAGTGAGCTGAGATGTTGCAATGTTGATGATATGCACTTCACTTGAAG 399
QY 133 ProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPhe 152
DB 400 CCATTTAGGGTTTTAGTTAGTATACAGGAGATTAACCCCTGAAACGCTACAGGCTCTTT 459
QY 153 AspAlaValGlnLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaValGlu 172
DB 460 GACGAGTCGAGAGACGATGCGGATTCGAATTTGATGATACATGTTTCTGATGAGTTAG 519
QY 173 ValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGlu 192
DB 520 GTTCAAGCTTTAGTGAACACAGGGTTGTTCTTCTTATGAAATGTTGATCAAGAG 579
QY 193 CysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaTrp 212
DB 580 TGTTCGGTGTGAGGAAAGTAGACCTTTGCGTCTTAAAGGCTCTTAAAGCTTGG 639
QY 213 IleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValValGln 232
DB 640 ATTACAGGACAGAGGAAAGCAATCTCCGGTACGAGATCTGAGATCCCTATTGTTGAG 699
```

```
QY 233 ValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsnPro 252
DB 700 GTTGATCCAGTGTTTGAAGGTTAGATGGCGGTGTTGGAAGTCTTGTGGAAGTGAATCCT 759
QY 253 ValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProVal 272
DB 760 TTGGCTAAATGTTGAAGGAGCTGATGTGTGAACCTTTCTGAGAACTATGATGTTCCGGTG 819
QY 273 AsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProVal 292
DB 820 AATGCATTGCACGACACAGGTAATGTTGTAATCGGGTGTGAGCGGTACTAGGCCCGTG 879
QY 293 LeuProGlyGlnHisGluArgGlyArgTyrTrpTrpGluAspAlaLysAlaLysGlu 312
DB 880 CTTCCAGGCCAACATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 939
QY 313 CysGlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluAspValAsnGlyAsnGly 332
DB 940 TGTGGTCTACACAAAGGGAACATC-----AAGGAGGAGAGAT----- 975
QY 333 LeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSerPro 352
DB 976 -----GGTCTCGACACTCAAAGCCCTGCTGTGCAAGAGATATTGAAAGCAAC 1026
QY 353 AsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArgLys 372
DB 1027 AATGTGTTGTCATTGAGCAAGAGGAGGTTGAGAACTTTTGAAGCTAGAGAACCGTAAA 1086
QY 373 GluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMetGluGlu 392
DB 1087 GAGCGCTGGTGTGCTGCTACTTTACGCTCTCTTGTGTCCTTTCTGCCAGGCGATGGAAGCA 1146
QY 393 SerTyrValAspLeuAlaAspLysLeuAlaGlySerThrClyMetLysValGlyLysPhe 412
DB 1147 TCGTACATCGAATTTGCTGAGAACTTCGCGGAAAA---GAGTTTAAAGTGGCGAAATTC 1203
QY 413 ArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPhePro 432
DB 1204 CGAGCTGACGGTGAGCAGAGAGAGAGTTTGTACCAAGAGCTTCAGTTAGGAGCTTCCCG 1263
QY 433 ThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSerGluLys 452
DB 1264 ACGATATCTTCTTTTCCGAAAGAGCTCCACGG---GCTATTAAAGTACCTTCAGAGCAT 1320
QY 453 ArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
DB 1321 AGAGATGTTGATTCATCTCATGCTGTTTGTGAATCTTCTTCGG 1362

RESULT 10
US-09-938-842A-2305
; Sequence 2305, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2305
; LENGTH: 1365
; TYPE: DNA
```

ORGANISM: Arabidopsis thaliana  
US-09-938-842A-2305

## Alignment Scores:

Pred. No.: 1.14e-189 Length: 1365  
Score: 1738.00 Matches: 346  
Percent Similarity: 82.5% Conservative: 45  
Best Local Similarity: 73.0% Mismatches: 55  
Query Match: 71.1% Indels: 28  
DB: 3 Gaps: 9

US-10-731-525-8 (1-466) x US-09-938-842A-2305 (1-1365)

```
Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThr---SerThrPhe----- 16
Db 1 ATGGCTTTAGCTGTTACTTCTTCAACTCGCAATCTCTGATCGAGATTTCTCGTTCT 60
Qy 17 ---ProSerSerGluProLysLeuProGlnIleGlySerIleArgIleSerGluArgPro 35
Db 61 GGAGCTTCTTCTGAATCTAAAGCTCTTCAATATGTTGATTTAGTTATCTGATCGAAC 120
Qy 36 IleGlyGlyAlaValAsnPheAsnLeuSerGlnArgSerLeuValLysProValAsn 55
Db 121 -----CATTTGCTCTCAGAGACGTTACTCTATGAAACCTTTAAAC 159
Qy 56 AlaGluPro---ProArgLysAspSerIleValProLeuAlaAlaThrIleValAla 74
Db 160 GCTGAGTCACATTCACGAGCAATCTTGGGTTACTCGTCTTCAACTCTAATTGCTCCT 219
Qy 75 SerAlaSerGluThr-----LysGluGluAspPheGluGlnIleAlaSerAspLeuAsp 92
Db 220 GAAGTTGAAGAGAAAGAGAGAGAGTTGAAGACTTTGAGCAACTTTGCTTAAAGCTTGAA 279
Qy 93 AsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAla 112
Db 280 GATGCTTCTCCACTTGAATCATGATAAGCTCTTCAGAGATTCGAGACCAATCGCA 339
Qy 113 IleAlaPheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArg 132
Db 340 ATTGCTTTTATGAGCTGAAGATGTTGCAATGATTGAATATGCACGTTTAACTGGAAG 399
Qy 133 ProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPhe 152
Db 400 CCATTTAGGGTTTATTAGTTAGATACAGGGAGATTAAACCTGAAAGCTACAGCTCTTT 459
Qy 153 AspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaValGlu 172
Db 460 GACGCACTCGAGAGACGATACGGGATTCGAATTCAGTACATGTTTCTGATGCAGTTGAG 519
Qy 173 ValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGlu 192
Db 520 GTTCAGCTTTAGTGAGGAACAAGGGTTGTTCTTCATTTATGAAGATGGTCAACAGAG 579
Qy 193 CysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaTyr 212
Db 580 TGTTCGCGTGTGAGAAAGTTAGACCTTTGCGTCTCTTAAGGGTCTTAAAGCTTGG 639
Qy 213 IleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValValGln 232
Db 640 ATTACAGACAGAGAAAGACCAATCTCCGGTACGAGATCTGAGATCCCTATTGTTTCAG 699
Qy 233 ValAspProAlaPheGluGlyMetAspGlyClyIleGlySerLeuValLysTyrAsnPro 252
Db 700 GTTTCATCAGTGTGTTGAAGGGTTAGATGGCGGTGTTGGAAGTCTTGTGAAGTGAATCCT 759
Qy 253 ValAlaAsnValLysGlyHisAspIleTyrAsnPheLeuArgThrMetAsnValProVal 272
Db 760 TTGGCTAATGTTGAAGGAGCTGATGTGGAACTTTCTGAGAACTATGGATGTTCCGGTG 819
Qy 273 AsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProVal 292
Db 820 AATGCAATTGCAGCACAAAGGGTATGTGTCAATCGGGTGTGAGCCGTGTACTAGCCCGTG 879
```

```
Qy 293 LeuProGlyGlnHisGluArgGluGlyArgTyrTyrIleGluAspAlaLysGlu 312
Db 880 CTTCCAGGCCAACATGAGAGAGAGAGAGTGTGTGTGGAGAGATGCTAAAGCTTAAGAA 939
Qy 313 CysGlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluAspValAsnGlyAsnGly 332
Db 940 TGTGGTCTTACACAAGGGAACATC-----AAGGAGGAAGAT----- 975
Qy 333 LeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSerPro 352
Db 976 -----GGTGTCTCAGACTCAAAAGCCTGCTGCTGCAAGAGATATTTGAAAGCAAC 1026
Qy 353 AsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArgLys 372
Db 1027 AATGTGGTTCATTTAGCAAGAGAGGGGTTGAGAACTTTTGAAGCTAGAGAACCGTAAA 1086
Qy 373 GluProTyrPheValValLeuTyrAlaProTyrCysProTyrCysGlnAlaMetGluGlu 392
Db 1087 GAGGCGTGGTGGTGTCTACTTTACGCTCTCTTGTGGCCCTTCTGCCAGGCGATGGAAGCA 1146
Qy 393 SerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLysPhe 412
Db 1147 TCGTACATCAAAATGGCTGAGAAACTTGGCGGAAAA---GGAGTTAAAGTGGCGAAATTC 1203
Qy 413 ArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPhePro 432
Db 1204 CGAGCTGACGGTGAAGCAAGAGAGTTGCTTAAGCAAGAGCTTCAGTTAGGGAGCTTCCCG 1263
Qy 433 ThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSerGluLys 452
Db 1264 ACATGACTTCTCTTCGAAAAAGAGCTCCACGG---GCTATTAAAGTACCCCTTCAGAGCAT 1320
Qy 453 ArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
Db 1321 AGAGATGTTGATTCACCTCATGCTGCTGTTGTGAATCTTCTTCGG 1362
```

## RESULT 11

US-10-731-525-9  
; Sequence 9, Application US/10731525  
; Publication No. US20040121440A1  
; GENERAL INFORMATION:  
; APPLICANT: Falco, Saverio Carl  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167-C  
; CURRENT APPLICATION NUMBER: US/10/731,525  
; CURRENT FILING DATE: 2003-12-09  
; PRIOR APPLICATION NUMBER: US/09/720,318A  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 1998-07-14  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 9  
; LENGTH: 1827  
; TYPE: DNA  
; ORGANISM: Triticum aestivum  
US-10-731-525-9

Alignment Scores:  
Pred. No.: 8.25e-179 Length: 1827  
Score: 1645.50 Matches: 327  
Percent Similarity: 79.2% Conservative: 49  
Best Local Similarity: 68.8% Mismatches: 78  
Query Match: 67.3% Indels: 21  
DB: 7 Gaps: 8

US-10-731-525-8 (1-466) x US-10-731-525-9 (1-1827)

```
Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSerGlu 20
Db 89 ATGGCTTCCGCT---ACTGCTTCCATCTCG-----TCGCACCTCCATCGCCCTCGCGAT 139
```



QY 21 ProLysLeuProGlnIleGlySerIleArg-----IleSerGluArgProIleGlyGly 38  
 DB 140 CTCAAGCCGCGAGATTGGAGCGGTGAGGCGAGCGAGCGGTGGTCTCTCGCGGCTG 199  
 QY 39 AlaValAsnPheAsnLeuSerGlnArgArgSerLeuValLysProValAsnAlaGluPro 58  
 DB 200 CCGCAACGGCGCCCAAGGCGCAGCGCGAGGCGGTGGCGCGGTGGCGGCGAG 259  
 QY 59 ProArgLysAspSerIleValProLeuAlaAlaThrIleValAla----- 74  
 DB 260 CCAGCGAGGAAGCCAGGTGCGGCTCCGCGGCTCTGTCGCGGTGGCGAGGAG 319  
 QY 75 SerAlaSerGluThrLysGluGluAspPheGluGlnIleAlaSerAspLeuAspAsnAla 94  
 DB 320 GAGGCATCTGCGTGGCGGCGGTGGAGTACGAGGCGCTGGCGAGGAGCTGGTGGGCGG 379  
 QY 95 SerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAla 114  
 DB 380 TCGCGCGTGGAGATCATGGATCGTGGCTCGACATGTTCCGCTCCGAAATCGGCATCGCC 439  
 QY 115 PheSerGlyAlaGluAspValAlaLeuIleGluTyrAlaLysLeuThrGlyArgProPhe 134  
 DB 440 TTCAGTGGTCCGAGGAGCGTGGCCCTCATCGAATACGCGAACTGACTGGACGCCCTTC 499  
 QY 135 ArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyrGlnLeuPheAspAla 154  
 DB 500 AGGGTGTTCAGCTTGCACCTGGCGGCTGAGTCCGAGTCCGAGGAGTCCGAGGAGTCCGAG 559  
 QY 155 ValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAspAlaValGluValGln 174  
 DB 560 GTGGAGAACGACATATGGATTCACATCGATGACATGTTCCCTGAGGCGCACGAGGTGCAA 619  
 QY 175 AlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAspGlyHisGlnGluCysCys 194  
 DB 620 GACCTTGTGAGGAGGAGGCGCTTTCAGTACGATGATGTTCTCTTTCAGGAGCGGACACGAGGTGCTGC 679  
 QY 195 ArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaIleThr 214  
 DB 680 AGGGTGAAGGAGTTCGCGCTTTCAGGAGGCGCTCAAGGCGCTCAAGGCGCTGGATCACC 739  
 QY 215 GlyGluArgLysAspGlnSerProGlyThrArgSerGluIleProValGlnValAsp 234  
 DB 740 GGGCAGAGGAAGATCATGTCCTGCGCACGAGGCGGATGCTGTTTCAAGTGTAT 799  
 QY 235 ProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsnProValAla 254  
 DB 800 CCGTCTTTGAAGGCTGGATGTGGAGCGGTAGTGTGATCAAGTGAACCTCTGGCT 859  
 QY 255 AsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAsnSer 274  
 DB 860 AATGTGGATGGCAAGGATATCTGGACCTTCTCAGGACCATGATGTCCTGTGAACACC 919  
 QY 275 LeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeuPro 294  
 DB 920 CTGCATGCTCAAGGCTACGCTCATGTTGGTGGAGCGGTGACACGCGCGGTGGTGGCG 979  
 QY 295 GlyGlnHisGluArgGluArgTyrTrpTrpGluAspAlaLysAlaLysGluCysGly 314  
 DB 980 GGGCAGCAGAGGAGGAGGAGGAGTGGTGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1039  
 QY 315 LeuHisLysGlyAsnValLys-----GlnGlnLysGluGluAspValAsnGlyAsn 331  
 DB 1040 CTGCACAAACCGTTAACTCGACAAAGGAAGGTTCAGGCGACCAAGGTTCGCGGTCAACGGCAAC 1099  
 QY 332 GlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIlePheAsnSer 351  
 DB 1100 GGCTCGCGCGAGCGGCGCGCT-----CCAGACATCTTCCAGAGC 1138  
 QY 352 ProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAspArg 371  
 DB 1139 CAGGCCATCTGTAACCTTCAACCGTCCGGGATCGAAGAACCTCTCGGCTCGAAGACCGC 1198  
 QY 372 LysGluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCysGlnAlaMetGlu 391

DB 1199 GCCGAGCGGTGGCTCAACCGTCTCTACGCTCCCTGGTGGCCATCTGCCAGGCAATGGAG 1258  
 QY 392 GluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLysValGlyLys 411  
 DB 1259 GCGTCTTACGTTGAGCTGGCGGAGAGCTGAGCGGCTCA---GGCATCAAGTGGCGCAAG 1315  
 QY 412 PheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPhe 431  
 DB 1316 TTCGCGCGGAGCGGAGGAGGAGCCATTCGCGAGGCGGAGCTGCAACTACAGAGCTTC 1375  
 QY 432 ProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSerGlu 451  
 DB 1376 CCGACGATCTCTGTTCCCGCGCGCAGCGCTGAAGCCC---ATCAAGTACCCGCTCGAG 1432  
 QY 452 LysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466  
 DB 1433 AAGAGGACGCTCCAGTCCCTCTCGGCTTCGTGAACAGCCTCAGA 1477

RESULT 12  
 US-10-762-049-9  
 : Sequence 9, Application US/10762049  
 : Publication NO. US20040139492A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Falco, Saverio Carl  
 : APPLICANT: Allen, Stephen M.  
 : TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
 : FILE REFERENCE: BB-1167-C  
 : CURRENT APPLICATION NUMBER: US/10762,049  
 : PRIOR FILING DATE: 2004-01-21  
 : PRIOR APPLICATION NUMBER: US/09/720,318A  
 : PRIOR FILING DATE: 2000-12-21  
 : PRIOR APPLICATION NUMBER: 60/092,833  
 : PRIOR FILING DATE: 1998-07-14  
 : NUMBER OF SEQ ID NOS: 12  
 : SOFTWARE: Microsoft Office 97  
 : SEQ ID NO 9  
 : LENGTH: 1827  
 : TYPE: DNA  
 : ORGANISM: Triticum aestivum  
 : US-10-762-049-9

Alignment Scores:  
 Pred. No.: 8,25e-179 Length: 1827  
 Score: 1645.50 Matches: 327  
 Percent Similarity: 79.2% Conservative: 49  
 Best Local Similarity: 68.8% Mismatches: 78  
 Query Match: 67.3% Indels: 21  
 DB: 7 Gaps: 8

US-10-731-525-8 (1-466) x US-10-762-049-9 (1-1827)

QY 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSerGlu 20  
 DB 89 ATGGCTTCGCTCT---ACTGCTTCATCTCG-----TCGCATCTCATCGCGCTGCGCAT 139  
 QY 21 ProLysLeuProGlnIleGlySerIleArg-----IleSerGluArgProIleGlyGly 38  
 DB 140 CTCAAAGCGCGGAGGATTGGAGCGGTGGAGCAGCAGGTGGCGGCTGTTCTTCGCGGCGCTG 199  
 QY 39 AlaValAsnPheAsnLeuSerGlnArgArgSerLeuValLysProValAsnAlaGluPro 58  
 DB 200 CCGGCAACGGCGCCCAAGGCGGCGGAGGCGGCGGCGGCTGCGCGCGGCGGAG 259  
 QY 59 ProArgLysAspSerIleValProLeuAlaAlaThrIleValAla----- 74  
 DB 260 CCAGCGAGGAAGCAGATGTCGCGCTCCGCGGCGCTCGTGGCGGCTGGCGGCGGAG 319  
 QY 75 SerAlaSerGluThrLysGluGluAspPheGluGlnIleAlaSerAspLeuAspAsnAla 94  
 DB 320 GAGGCATCTGCGGCTGGCGGCGGTGGAGCTACGAGGCGCTCGGCGGAGGCGGTGGTGGGCGCG 379  
 QY 95 SerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAlaIleAla 114



```
Db 380 TCAGCGCTGGAGATCATGGATGGCGCTGCATGTTGGCTCCGAAATCGCATCGCC 439
Qy 115 PheSerGlyAlaGluAaspValAlaLeuIleGluTyrAlaLysLeuThrGlyArgProPhe 134
Db 440 TTCAGTGGTCCGAGGACGTGGCCCTCATCGAATACGGAACCTGACTGGACGCCCTTC 499
Qy 135 ArgValPheSerLeuAaspThrGlyArgLeuAaspProGluThrTyrGlnLeuPheAaspAla 154
Db 500 AGGTGTTCAGCTTGACACTGGCGGCACTGAACCCAGAGACATACGAACCTTCGCAAG 559
Qy 155 ValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAaspAlaValGluValGln 174
Db 560 GTGGAGAAGCACTATGATCCATCGATGATACATGTTCCCTGAGGCCAGGAGTGCAG 619
Qy 175 AlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAaspGlyHisGlnGluCysCys 194
Db 620 GACCTTGTGAGGACCAAGGGCTCTTCTCTTCTACGAGGACGACACAGGAGTGTGTC 679
Qy 195 ArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaTrrIleThr 214
Db 680 AGGTGAGGAAGTTCGGCCCTTGAGGAGGCCCTCAAGGGCCCTCAAGGCCCTGGATCACC 739
Qy 215 GlyGlnArgLysAaspGlnSerProGlyThrArgSerGluIleProValGlnValAasp 234
Db 740 GGGCAGAGAAGGATCAGTCCCTGGCACCCAGCCAGCATCCCTGTGTTCAGTTGAT 799
Qy 235 ProAlaPheGluGlyMetAaspGlyGlyIleGlySerLeuValLysTrrAaspProValAla 254
Db 800 CCGTCTTTGAAGGGCTGGATGGTGGAGCCGCTAGCTTGATCAAGTCAAGCCCTGTGGCT 859
Qy 255 AsnValLysGlyHisAaspIleTrrAaspPheLeuArgThrMetAsnValProValAaspSer 274
Db 860 AATGTGATGCAAGGATATCTGACCTTCTCAGGACCATGATGTCCTGTGTGAACACC 919
Qy 275 LeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeuPro 294
Db 920 CTGATGCTCAAGGCTAGCTCTCATTTGGGTGGAGCGGTGCACAGGCCGCTGTGGCG 979
Qy 295 GlyGlnHisGluArgGluGlyArgTrrPrrPrrGluAaspAlaLysAlaLysGluCysGly 314
Db 980 GGGCAGCAGCAGAGAGGAGGAGGTGGTGGTGGAGGACGCCACGGCCAGGAGTGGCGC 1039
Qy 315 LeuHisLysGlyAsnValLys-----GlnGlnLysGluGluAaspValAaspGlyAsn 331
Db 1040 CTGCACAAACGGTAATATCGACAAGAAAGGTGAGGACCCCAAGGTGGCGCTCAACGGCAAC 1099
Qy 332 GlyLeuSerGlnSerHisAlaAsnGlyAaspAlaThrValProAaspIlePheAsnSer 351
Db 1100 GGCTCGGCGAGGCCAGCGCC-----CCAGACATCTTCCAGAGC 1138
Qy 352 ProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeuGluAaspArg 371
Db 1139 CAGGCCATCGTCAACCTCACCCGTCGCGGATCGAGAACCTCTCTGCGGCTCGAGAACCGC 1198
Qy 372 LysGluProTrrPrrLeuValLeuTyrAlaProTrrCysProTrrCysGlnAlaMetGlu 391
Db 1199 GCCAGGCCGTGGCTCACCGTCTCTACGCTCCCTGGTGCCCATCTCGCAGGCAATGGAG 1258
Qy 392 GluSerTyrValAaspLeuAlaAaspLysLeuAlaGlySerThrGlyMetLysValGlyLys 411
Db 1259 CGGTCTCTAGTGTAGTGGCCGAGAGAGTTCAGCGGCTCA---GGCATCAAGGTGGCCAG 1315
Qy 412 PheArgAlaAaspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeuGlySerPhe 431
Db 1316 TTCGCGCGGACGGCGCAGCAGAACCAATTCGCGAGCGGAGCTGCAACTACAGAGTTC 1375
Qy 432 ProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyrProSerGlu 451
Db 1376 CCGACGATCTCTCTGTTCCCGGCGCACCGTGAAGCCC---ATCAAGTACCCCGTCCGAG 1432
Qy 452 LysArgAaspValAaspSerLeuMetAlaPheValAsnAlaLeuArg 466
```

```
Db 1433 AAGAGGACGCTCCAGTCCCTCTCTCTCGCTTCTGTAACAGCCTCAGA 1477
RESULT 13
US-10-739-930-2784
; Sequence 2784, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 2784
; LENGTH: 1840
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1840)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: ZEAMA-23APR03-CLUSTER4907_2
US-10-739-930-2784
Alignment Scores:
Pred. No.: 8,34e-179 Length: 1840
Score: 1645.50 Matches: 329
Percent Similarity: 78.2% Conservative: 47
Best Local Similarity: 68.4% Mismatches: 67
Query Match: 67.3% Indels: 39
DB: 8 Gaps: 10
US-10-731-525-8 (1-466) x US-10-739-930-2784 (1-1840)
Qy 1 MetAlaLeuAlaPheThrSerSerIleSerAlaProThrSerThrPheProSerSerGlu 20
Db 114 ATGGCATCCGTACCACCTCCATCTCGTCGCACCTCGCCGCTC-----CGCGAC 164
Qy 21 ProLysLeuProGlnIleGlySerIleArgIleSerGluArgProIleGlyAlaVal 40
Db 165 GTCAAGCTCGAGGATTTGGTGC-----GCGCTGAGGAGGTGGCGCTGGCGCGC 215
Qy 41 AsnPheAsnLeuSerGlnArgSerLeuValLysProValAen---AlaGluProPro 59
Db 216 CGT-----GGCCAGCGCTCGGGCGGTGCCCCGCTGCGCGCGCGGACCTCGC 266
Qy 60 ArgLysAaspSerIleValProLeuAlaAlaThrThrIleValAlaSer-----AlaSer 77
Db 267 AGGCGG-----CCGGTGGCGGCTTCCGCGNCCGCGGTGC-CCCNCGCGCGCG 313
Qy 78 GluThrLysGluGlu-----AaspPheGluGlnIleAlaSerAaspLeuAaspAla 94
Db 314 CCCGTGCGCAGCAGGCGGCGGTGGACTACGAGGCCCTGGCGCGCGAGCTTGAGGCGCG 373
Qy 95 SerProLeuGluIleMetAaspArgAlaLeuAaspLysPheGlyAsnAaspIleAla 114
Db 374 TCGCCGTGGAGATCATGGATTCGGCGGTGCGCCATGTTCCGAATCCGAATCGCATGCC 433
Qy 115 PheSerGlyAlaGluAaspValAlaLeuIleGluTyrAlaLysLeuThrGlyArgProPhe 134
Db 434 TTCAGCGCGCGGAGGACGTGGCGCTGATCGAGTACGCGAAGCTGACGGGCGTCCCTTC 493
Qy 135 ArgValPheSerLeuAaspThrGlyArgLeuAaspProGluThrTyrGlnLeuPheAaspAla 154
Db 494 CGGCTGTTACGCTGGACACCGGGCGGCTCAACCGGAGACGTACCACTCTTCGACAGG 553
Qy 155 ValGluLysHisTyrGlyIleArgIleGluTyrMetPheProAaspAlaValGluValGln 174
Db 554 GTGAGAAGCACTACGGCATCCGATCAGTACATGTTCCCGGACGCGAGGAGTGCAG 613
Qy 175 AlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAaspGlyHisGlnGluCysCys 194
```

[illegible]

RESULT 14  
US-10-437-963-23751  
; Sequence 23751, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovacic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei

```

; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)/B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 23751
; LENGTH: 1880
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_28800C.1
US-10-437-963-23751

Alignment Scores:
Pred. No.: 4,24e-178 Length: 1880
Score: 1639.50 Matches: 324
Percent Similarity: 78.3% Conservative: 51
Best Local Similarity: 67.6% Mismatches: 71
Query Match: 67.1% Indels: 34
DB: Gaps: 10

US-10-731-525-8 (1-466) x US-10-437-963-23751 (1-1880)

QY 10 SerAlaProThrSerThrPheProSerSerGluProLysLeuProGlnIleGlySerIle 29
DB 119 TCTCGTCCCACTCCGTGC-----CCCTCGCGACCTCAAGCTGCG-AGGATTGGAGCGCGTG 174
QY 30 Arg-----IleSerGluArgProIleGlyGlyAlaValAsnPheAsnLeuSerGlnArg 47
DB 175 AAGCAGCAGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 231
QY 48 ArgSerLeuValLysProValAsnAlaGluProProArgLysAspSerIleValProLeu 67
DB 232 GCAGAGGCGCGTGCCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 282
QY 68 AlaAlaThrThrIleValAlaSerAlaSerGluThrLysGluGlu----- 82
DB 283 TCGCGCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 342
QY 83 -----AspPheGluGlnIleAlaSerAspLeuAspAsn 93
DB 343 GCGGTGACGCCCTTGCTCCGCGCGTGGACTACGAGGCGCTGGCGCCAGGAGTGCAGGCGC 402
QY 94 AlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGlyAsnAspIleAlaIle 113
DB 403 GCGTCGCGCGTGGAGATCATGGATCGGCGCACTCGCCATGTTCCGATCCGACATCGCCATC 462
QY 114 AlaPheSerGlyAlaGluAspValAlaLeuIleGluThrAlaLysLeuThrGlyArgPro 133
DB 463 GCCTTCAGCGGGGCGGAGACGTGGCGCTGATCAGATACGCGAAGCTACGCGGGCGTCCG 522
QY 134 PheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThrTyxGlnLeuPheAsp 153
DB 523 TTCGCGGTGTTACGCTTGGACACGGGCGCGGTGAACCGGAGACGTACAGGTGTTCCGAC 582
QY 154 AlaValGluLysHisThrGlyIleArgIleGluThrMetPheProAspAlaValGluVal 173
DB 583 AAGGTGGAGAAGCACACTACGGGATCCGGATCGAGTACATGTTCCCGGACGCCCGCGAGGTG 642
QY 174 GlnAlaLeuValArgSerLysGlyLeuPheSerPheThrGluAspGlyHisGlnGluCys 193
DB 643 CAGCGCGTGTGGCGCGCAAGGGGCTCTTCTCTTCTACGAGGACGCCCGCAGGAGTGC 702
QY 194 CysArgValArgLysValArgProLeuArgArgAlaLeuLysGlyLeuArgAlaTrpIle 213
DB 703 TGCGCGCGCCGCAAGGTGCGCGCGCTTCGCGCGCGCGCTCCGCGGCTCAGGCGCTGGAATC 762
QY 214 ThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluIleProValGlnVal 233

```

```
Db 763 ACGGGCAGCGCAAGACCAAGTCCCGCGGACCGCCGCGCCATCCCGCTCGTCCAGGTC 822
Qy 234 AspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuValLysTrpAsnProVal 253
Db 823 GACCCCTCTTCGAGGCGCTCGTGGCGGCGCGCGAGCTTCGTCAGAGTGAACCCCGTC 882
Qy 254 AlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMetAsnValProValAsn 273
Db 883 GCCACGCTCGACGCGCAGGAGCTGTGACCTTCCTCCGCCCATGGATGTCCTCGTCAAC 942
Qy 274 SerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCysThrArgProValLeu 293
Db 943 GCCCTGACGCGCCAGGCGTACGTCGATCGGTCGAGCGCTGCACAGCGCCCGTCTT 1002
Qy 294 ProGlyGlnHisGluArgGlyArgTrpTrpTrpGluAspAlaLysGluCys 313
Db 1003 CCCGGCCAGCAGCAGCGGGAAGGCGGTGGTGGGAGACGC CAAGGCCAAGAGTGC 1062
Qy 314 GlyLeuHisGlyAsnValLysGlnGlnLysGluAspValAsnGlyAsnGlyLeu 333
Db 1063 GGCTCCACAGGCGACATC-----GACGACCGAGGTGGGGCGCGCG 1107
Qy 334 SerGlnSerHis-----AlaAsnGlyAspAlaThrThr-----ValProAspIle 348
Db 1108 CGCGCGCGCACAAAGCGCGCGGCCCAACGCCAACCGGTCCGCGCGCGCGCCGACATC 1167
Qy 349 PheAsnSerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeu 368
Db 1168 TTCGAGAGCAGCGCGGTGGTGGCTCACC CGCGCGCGGTGCGAGAACCTGCTCCGCGTG 1227
Qy 369 GluAspArgLysGluProTrpLeuValValLeuTyrAlaProTrpCysProTyrCysGln 388
Db 1228 GAGAGCGCGCGAGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1287
Qy 389 AlaMetGluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySer---ThrGlyMet 407
Db 1288 GCCATGAGGCGTCTACCTGAGCTCGCGAGAGGCTGGCGCGCGCGCGCGCGCGCGGTG 1347
Qy 408 LysValGlyLysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGln 427
Db 1348 AAGTGGGCAAGTTCGCGCGCGAGCGGCGAGCAGAAAGCGGTTCGCGCAGAGGAGCTGCAG 1407
Qy 428 LeuGlySerPheProThrIleLeuPhePheProLysHisSerSerArgProThrIleLys 447
Db 1408 CTGACAGGCTTCGCGACCATCTCTCTCCGAGCGCCACCGGAGGCGG---ATCAAG 1464
Qy 448 TyrProSerGluLysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
Db 1465 TACCCGTCGGAGAGCGGAGCTGCAGCTCGCTCTCGCTCTCGTCAACAGCCTCCGA 1521
```

## RESULT 15

```
US-10-169-667A-1
; Sequence 1, Application US/10169667A
; Publication No. US2004012341A1
; GENERAL INFORMATION:
; APPLICANT: BIOEMVA
; TITLE OF INVENTION: Method for obtaining plants enriched in cysteine and glutathione
; FILE REFERENCE: BET 02/0583
; CURRENT APPLICATION NUMBER: US/10/169,667A
; CURRENT FILING DATE: 2003-03-28
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1380
; TYPE: DNA
; ORGANISM: Lemna minor
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1377)
US-10-169-667A-1
Alignment Scores:
Pred. No.: 1.89e-173 Length: 1380
```

```
Score: 1597.50 Matches: 315
Percent Similarity: 78.0% Conservative: 58
Best Local Similarity: 65.9% Mismatches: 66
Query Match: 65.3% Indels: 39
DB: 7 Gaps: 11

US-10-731-525-8 (1-466) x US-10-169-667A-1 (1-1380)
Qy 9 IleSerAlaProThrSerThrPheProSerSerGluProLysLeuProGlnIleGlySer 28
Db 1 ATGTGGCGCCACAGCAGCAACTATGGCGGCTCG-----CtCTCTCA 42
Qy 29 IleArgIleSerGluArgProIleGlyGlyAlaValAsnPheAsnLeuSerGlnArg 48
Db 43 CACAGCGTCTCTCCCGCGAATTTTCAGGAGGATTCAGAGTTCAGCTCTACGAGGTGAGG 102
Qy 49 SerLeu-----ValLysProValAsnAla---GluProProArg 60
Db 103 AAATTGACAGAGGAGCGGTTTGTCTCCGATCGCGCTCTGTGTGCGGTGGATCCAGGAGG 162
Qy 61 LysAspSer-----IleValProLeuAla-----AlaThr 70
Db 163 AGAGCCAGTCGCGAGTGGCGCGCTGGCGCGCGTCCCTCCACCTCTGATGCCGTG 222
Qy 71 ThrIleValAlaSerAlaSerGluThrLysGluGlu-----AspPheGluGlnIleAla 88
Db 223 ACTCTGTCTCCAGCAGCAGAAACGAGCAGCGCTTGATTTGATTTGAGAAGCTCGCG 282
Qy 89 SerAspLeuAspAsnAlaSerProLeuGluIleMetAspArgAlaLeuAspLysPheGly 108
Db 283 AGCGAGCTAGAGCGCGCATCGCTCTCGAGATTATGACAAAGGCTCTGGAGATGTTCCGT 342
Qy 109 AsnAspIleAlaIleAlaPheSerGlyValaGluAspValAlaLeuIleGluTyrAlaLys 128
Db 343 AACGATATTCATCGCTTCAGCGGCGCGAGAGCGTTCGTCTGATCGAGTACGACGC 402
Qy 129 LeuThrGlyArgProPheArgValPheSerLeuAspThrGlyArgLeuAsnProGluThr 148
Db 403 CTAACCGAGCAGCTTTTCAGGTTTCAGTCTCGACACCGCAGCGCTAAACCCAGAGACA 462
Qy 149 TyrGlnLeuPheAspAlaValGluLysHisTyrGlyIleArgIleGluTyrMetPhePro 168
Db 463 TACCGATTTCTCGATGAAGTAGAGAGCGGTACAACTCCGATCGAGTACATGTTCCCC 522
Qy 169 AspAlaValGluValGlnAlaLeuValArgSerLysGlyLeuPheSerPheTyrGluAsp 188
Db 523 GACCGGTGGAGTCCAGGCCCTCGTCGAAGCAAGGCCCTCTTTCTCTCTACGAGGAC 582
Qy 189 GlyHisGlnGluCysCysArgValArgLysValArgProLeuArgArgAlaLeuLysGly 208
Db 583 GGGCACCAGGAGTGTCTGCAGAGTGAGAAAGGTGAGGCCCTTTAGAGAGCTCTCAGAGGC 642
Qy 209 LeuArgAlaTrpIleThrGlyGlnArgLysAspGlnSerProGlyThrArgSerGluTrle 228
Db 643 CTCAGAGCTTGATTCACCGCGCAGCGCAAGACAGTCTCCCGCACCCGAGCAGCGGTT 702
Qy 229 ProValValGlnValAspProAlaPheGluGlyMetAspGlyGlyIleGlySerLeuVal 248
Db 703 CCCACGTCAGGTGGATCTCTCGTTGAGGGATTTGAAGCGGACACGAGGAGCTTAATC 762
Qy 249 LysTrpAsnProValAlaAsnValLysGlyHisAspIleTrpAsnPheLeuArgThrMet 268
Db 763 AAATGGAACCCCGTGGCGAACGTAGACGGGCGAGATATCTGGAGATTTCTGAGAACGATG 822
Qy 269 AsnValProValAsnSerLeuHisAlaLysGlyTyrValSerIleGlyCysGluProCys 288
Db 823 GCCGTTCAGTGAATCTCCCTCACTCCCAAGGGTACGTGTCAATCGGTGCGAGCCATGC 882
Qy 289 ThrArgProValLeuProGlyGlnHisGluArgGluGlyArgTGTTrpTrpGluAspAla 308
Db 883 ACGCGCGCGTCTTACCCAGGACACACGAGAGGAGGAGAGGTGGTGGTGGGAGGACCGG 942
Qy 309 LysAlaLysGluCysGlyLeuHisLysGlyAsnValLysGlnGlnLysGluGluAspVal 328
```

```
Db 943 AAAGCCAAAGGAATCGGGCTCCACAAGGGCAACATCAAGCAA-----GACGAGTCCGCTG 996
Qy 329 AsnGlyAsnGlyLeuSerGlnSerHisAlaAsnGlyAspAlaThrThrValProAspIle 348
Db 997 -----CCTCCCTCCAGCAACGGCAACGGCACCGCGCGCTCGGATCTC 1038
Qy 349 PheAsnSerProAsnValValAsnLeuSerArgThrGlyIleGluAsnLeuAlaLysLeu 368
Db 1039 TTC-----TCCGTCGTTTCTCTGAGCGCGCCAGGSCATGGAGAACTTGCTCCGCTG 1089
Qy 369 GluAspArgLysGluProThrLeuValValLeuTyrAlaProThrCysProTyrCysGln 388
Db 1090 GAGAGCCGGAGGACTCGTGGCTTGTCTGCTCTATGCTCCTCGTGGCTCTCTGCGCAG 1149
Qy 389 AlaMetGluGluSerTyrValAspLeuAlaAspLysLeuAlaGlySerThrGlyMetLys 408
Db 1150 GGGATGGAGGAGTCGTTCAACGAGGTGGCAGGACTACTCGGCGCGGACTCAGGAGTGAAG 1209
Qy 409 ValGlyLysPheArgAlaAspGlyGluGlnLysGluPheAlaLysSerGluLeuGlnLeu 428
Db 1210 GTGGGCAAAATTCAGACCGGAGACCGGAGAGCTTTGCGGAG---GAGATCGGGCTG 1266
Qy 429 GlySerPheProThrIleLeuPhePheProLysHisSerSerArgProThrIleLysTyr 448
Db 1267 GTGAGTTTTCACCATACTCTTCTCCGAAAGCACGCTCGAAACCG---GTCAAGTAC 1323
Qy 449 ProSerGluLysArgAspValAspSerLeuMetAlaPheValAsnAlaLeuArg 466
Db 1324 CCGTCGGAGCAGAGAGACGTGACTCTCTGCTCGCCCTTCGTCAACGCCCTTCGC 1377
```

Search completed: February 20, 2006, 23:30:28

Job time : 1425 secs

**THIS PAGE LEFT BLANK**